ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

ON THE NON-EXISTENCE OF REGULAR STATIONARY SOLUTIONS OF RELATIVISTIC FIELD EQUATIONS

BY A. EINSTEIN AND W. PAULI

(Received January 4, 1943)

It is shown that the field equations of the relativistic gravitational theory and of its five-dimensional generalization do not admit any non-singular stationary solution which represents a field of non-vanishing total mass or charge.

Introduction

Some time ago one of us proved[1] that there exist no solutions of the gravitational equations $R_{ik} = 0$ satisfying the following conditions:

1) The field is stationary (i.e. the g_{ik} are independent of x^4).

2) It is free from singularities.

3) It is imbedded in a Euclidean space (of the Minkowski type), and for large values of r (r being the distance from the origin of the spatial coordinate system) g_{44} has the asymptotic form

$$g_{44} = -1 + \frac{\mu}{r},$$

where $\mu \neq 0$.

The third condition implies that the total gravitational mass of the field is different from zero.

The following considerations led us to reanalyze this proof, to reduce it to its necessary elements, and to generalize it to cases of higher dimensions.

When one tries to find a unified theory of the gravitational and electromagnetic fields, he cannot help feeling that there is some truth in *Kaluza's* five-dimensional theory. Yet its foundation is unsatisfactory in so far as, with respect to the group of admissible coordinate transformations, the fifth, space-like, coordinate is treated quite differently from the others. Consequently the components of the electromagnetic field transform independently from those of the gravitational field, and the two fields are only apparently unified. The question arises whether one could base the theory on the full group of five-dimensional point transformations without sacrificing its main achievements.

This might seem impossible, for according to all our experience the physical continuum has $3 + 1$ but not $4 + 1$ dimensions, since its objects appear to have three, but not four spatial dimensions. One could, however, imagine that this difficulty might be overcome as follows. Assume that in such a theory the fields corresponding to non-singular solutions are not point-like but linearly extended in a four-dimensional space. The geometrical configuration of several coexisting fields of this character would, then, more or less resemble the configuration of the objects of a three-dimensional space.

[1] A. Einstein, Revista (Universidad Nacional de Tucuman) A, **2**, 11, 1941.

We therefore have to investigate the question whether, in a five-dimensional metric continuum (of signature 1) the equations

$$R_{ik} = 0$$

admit of non-singular stationary solutions with a field g_{ik} asymptotically given by

$$\begin{matrix} A & 0 & 0 & 0 & 0 \\ 0 & A & 0 & 0 & 0 \\ 0 & 0 & A & 0 & 0 \\ 0 & 0 & 0 & B & D \\ 0 & 0 & 0 & D & C \end{matrix}$$

where at least one of the quantities, A, B, C, D has the form

$$\pm 1 + \frac{\text{const}}{r}$$

with a non-vanishing constant. This is the asymptotic form of a field representing a particle whose electric and ponderable masses do not both vanish.

By discussing the field equations, V. Bargmann[2] has shown that *spherically symmetric* solutions of this character do not exist. In what follows we shall prove the non-existence of regular solutions of the required form irrespective of any assumptions about the symmetry of the field (in regions of finite field strength).

The proof makes it clear why one always encounters singularities if one attempts to represent material particles by solutions of field equations which are based on Riemann's tensor.

1. According to Palatini the variation of the contracted Riemannian curvature tensor[3]

$$R_{IK} = \Gamma^S_{IS,K} - \Gamma^S_{IK,S} + \Gamma^R_{SI}\Gamma^S_{RK} - \Gamma^R_{IK}\Gamma^S_{RS} \tag{1}$$

can be written as

$$\delta R_{IK} = (\delta\Gamma^S_{IS})_{;K} - (\delta\Gamma^S_{IK})_{;S}\,, \tag{2}$$

or

$$\delta R_{IK} = U^S_{IK;S} \tag{3}$$

with

$$U^S_{IK} = -\delta\Gamma^S_{IK} + \tfrac{1}{2}(\delta\Gamma^A_{IA}\delta^S_K + \delta\Gamma^A_{KA}\delta^S_I). \tag{4}$$

This leads to

---

[2] Private communication.

[3] Capital indices assume all values 1, $\cdots$, n, n being the dimension of the space which we consider. Ordinary differentiation is denoted by a comma, covariant differentiation by a semicolon.

$$\sqrt{|g|}\, g^{IK}\delta R_{IK} = \sqrt{|g|}\, g^{IK} U^{S}_{IK;S}\,, \tag{5}$$

where $|g|$ is the absolute value of the determinant of the covariant metric tensor. Since the covariant derivatives of the metric tensor and of its density vanish, we can put the right side into the form

$$(\sqrt{|g|}\, g^{IK} U^{S}_{IK})_{;S}$$

and, finally, replace covariant by ordinary differentiation, the expression within the brackets being a contravariant vector density. Hence we get

$$\sqrt{|g|}\, g^{IK}\delta R_{IK} = -\mathfrak{A}^{S}{}_{,S} \tag{6}$$

with

$$\mathfrak{A}^{S} = \sqrt{|g|}\, (g^{IK}\delta\Gamma^{S}_{IK} - g^{AS}\delta\Gamma^{B}_{AB}). \tag{7}$$

If both the original and the varied fields satisfy the gravitational equations $R_{IK} = 0$, the δR_{IK} in (6) vanish, and we have

$$\mathfrak{A}^{S}{}_{,S} = 0. \tag{6a}$$

This is always the case when the variation of the field is produced by an infinitesimal change of the coordinate system. In performing such a variation, we must not compare the values of Γ or R at the same world point, but we have to *displace* the world point so as just to compensate the variation of its coordinates due to the change of the coordinate system. Only then can variation and differentiation be interchanged so that we have, for example,

$$(\delta\Gamma^{S}_{IK})_{,L} = \delta(\Gamma^{S}_{IK,L}).$$

For a variation produced by the coordinate transformation

$$x'^{I} = x^{I} + \xi^{I}(x)$$

we then have, retaining only terms of the first order in the ξ^{I},

$$\delta\Gamma^{L}_{IK} = -\xi^{L}{}_{,IK} + \xi^{L}{}_{,R}\Gamma^{R}_{IK} - \xi^{R}{}_{,I}\Gamma^{L}_{RK} - \xi^{R}{}_{,K}\Gamma^{L}_{RI} - \xi^{R}\Gamma^{L}_{IK,R}\,. \tag{8}$$

When (8) is inserted in (7), (6a) must hold for arbitrary functions ξ^{I} in consequence of the field equations $R_{IK} = 0$.

2. We now decompose the n-dimensional continuum of the x^{I} into the ordinary three-dimensional space of the x^{I} (lower case italic indices = 1, 2, 3) and the subspace of the remaining coordinates x^{ρ} (Greek indices 4, $\cdots$, n). We expressly assume the g_{IK} to be independent of the coordinates x^{ρ}:

$$\frac{\partial g_{IK}}{\partial x^{\rho}} = 0. \tag{9}$$

In the physical interpretation of this formalism, x^{4} is the time coordinate and x^{5} the fifth coordinate introduced in *Kaluza's* theory (where the metric is space-

like with respect to this fifth coordinate). For the following mathematical discussions, however, it is more convenient to restrict neither the dimension of the continuum nor the signature of its metric except for the assumption that the metric of the three-dimensional space be positive definite.

We are going to discuss the implications of the relations (6a) for those variations (8) which do not affect the condition (9) (cylindricity condition). Only the terms with $S = i$ have then to be retained in (6a). Using Gauss' theorem, we infer from $\mathfrak{A}^i{}_{,i} = 0$ that

$$\oint \mathfrak{A}^i n_i \, df = 0 \tag{10}$$

if the integral is extended over a closed two-dimensional surface which does not enclose any singularities of the field. n_i are the covariant components of a unit vector normal to this surface. In case singularities of the metric tensor do exist, we consider two different closed surfaces F_1 and F_2 as inner and outer boundaries of a three-dimensional region free from singularities. Gauss' theorem then leads to

$$\oint_{F_1} \mathfrak{A}^i n_i \, df = \oint_{F_2} \mathfrak{A}^i n_i \, df. \tag{10a}$$

There are two different types of infinitesimal coordinate transformations which leave the stationary character of the field [cf. (9)] invariant. The *first type* is characterized by functions ξ^I which are *independent* of the x^ρ but may depend in an arbitrary way on the x^i. In this case we may choose the ξ^I, together with their first and second derivatives, equal to zero on the inner surface F_1 so that the $\mathfrak{A}^i$ too disappear on F_1. Hence for this type of variations the stronger relation (10) holds even when the surface encloses singularities of the field, provided the surface itself is free from singularities.

The *second type* of transformations not affecting the stationary character of the field leads to an integral theorem which singles out the *regular* solutions of the field equations $R_{IK} = 0$. These transformations are defined by

$$\xi^i = 0, \qquad \xi^\rho = c^\rho_\sigma x^\sigma \tag{11}$$

with constant coefficients c^ρ_σ. From (8) and (9) it then follows that

$$\left.\begin{aligned} &\delta\Gamma^l_{ik} = 0, \quad \delta\Gamma^B_{iB} = \delta\Gamma^\rho_{i\rho} = 0, \quad \delta\Gamma^B_{\rho B} = \Gamma^B_{\rho B} = 0 \\ &\delta\Gamma^i_{\rho\sigma} = -(c^\tau_\rho \Gamma^i_{\tau\sigma} + c^\tau_\sigma \Gamma^i_{\rho\tau}), \qquad \delta\Gamma^l_{i\rho} = -c^\tau_\rho \Gamma^l_{i\tau} \end{aligned}\right\}. \tag{12}$$

Inserting these expressions in (7), we get

$$\mathfrak{A}^i = -2c^\tau_\rho \sqrt{|g|}\, g^{\rho A} \Gamma^i_{A\tau}.$$

Since the coefficients c^τ_ρ may be chosen arbitrarily, (6a) implies, in consequence of the field equations and of the stationary character of the solution,

$$(\sqrt{|g|}\, g^{\rho A}\Gamma^i_{A\sigma})_{,i} = 0 \tag{13}$$

for all values of ρ and σ. For *non-singular* solutions this leads to

$$\oint \sqrt{|g|}\, g^{\rho A}\Gamma^i_{A\sigma} n_i\, df = 0 \tag{13a}$$

in analogy with (10), while singular solutions satisfy only a weaker condition of the type (10a).

3. To make use of this theorem, we introduce the assumption that asymptotically the g_{IK}-field in question approaches that of Euclidean space. Hence for large distances from the origin of the coordinate system we may put

$$g_{IK} = \mathring{g}_{IK} + \gamma_{IK}\,. \tag{14}$$

Here the constants $\mathring{g}_{IK}$ can be brought to the form $\delta_{IK}\epsilon_I$ (no summation!) with $\epsilon_i = +1$, $\epsilon_\rho = \pm 1$; γ_{IK} are considered small of the first order. The determinant g is then given by

$$g = (\prod_I \epsilon_I)\cdot(1 + \gamma) \tag{15}$$

with

$$\gamma = \mathring{g}^{IK}\gamma_{IK} = \sum_I \epsilon_I \gamma_{II}\,. \tag{15a}$$

Retaining only first order terms, we infer from (1), because of the stationary character of the field, that the equations $R_{IK} = 0$ assume the form

$$\gamma_{IK,ss} - \gamma_{Is,sK} - \gamma_{Ks,sI} + \gamma_{,IK} = 0 \tag{16}$$

or

$$\left.\begin{aligned} &\gamma_{ik,ss} - \gamma_{is,sk} - \gamma_{ks,si} + \gamma_{,ik} = 0 \\ &\gamma_{i\rho,ss} - \gamma_{\rho s,si} = 0 \qquad \gamma_{\rho\sigma,ss} = 0 \end{aligned}\right\}. \tag{16a}$$

As is well known, it is always possible to normalize the coordinate system in such a way that

$$\gamma_{Is,s} - \tfrac{1}{2}\gamma_{,I} = 0, \tag{17}$$

or

$$\gamma_{is,s} - \tfrac{1}{2}\gamma_{,i} = 0 \qquad \gamma_{\rho s,s} = 0 \tag{17a}$$

then (16) reduces to Laplace's equation

$$\gamma_{IK,ss} = 0. \tag{18}$$

In what follows we consider only those terms which, at infinity, do not decrease more rapidly than r^{-1} (where $r^2 = \sum_i (x^i)^2$). We have, therefore,

$$\gamma_{IK} = m_{IK}/r. \tag{19}$$

Since

$$\gamma = \left(\sum_I \epsilon_I m_{II}\right) \Big| r = \left(\sum_i m_{ii} + \sum_\rho \epsilon_\rho m_{\rho\rho}\right) \Big| r,$$

we infer from (17a)

$$m_{ik} = \tfrac{1}{2}\delta_{ik}\left(\sum_j m_{jj} + \sum_\rho \epsilon_\rho m_{\rho\rho}\right), \qquad m_{i\rho} = 0,$$

or

$$m_{i\rho} = 0, \qquad m_{ik} = m\delta_{ik}, \qquad m + \sum_\rho \epsilon_\rho m_{\rho\rho} = 0. \tag{20}$$

Hence, neglecting terms of order higher than the first, we have

$$g_{ik} = \delta_{ik}\left(1 + \frac{m}{r}\right), \qquad g_{i\rho} = 0, \qquad g_{\rho\sigma} = \epsilon_\rho \delta_{\rho\sigma} + \frac{m_{\rho\sigma}}{r}, \tag{21}$$

where m satisfies the last equation (20). This result may be used to evaluate the integral (13a) for a sufficiently large sphere. Since on its surface we may put

$$\Gamma^i_{\rho\sigma} = \Gamma_{i,\rho\sigma} = -\tfrac{1}{2}\gamma_{\rho\sigma,i} = -\tfrac{1}{2}m_{\rho\sigma}\left(\frac{1}{r}\right)_{,i} = \tfrac{1}{2}m_{\rho\sigma}\frac{x^i}{r^3} = \tfrac{1}{2}m_{\rho\sigma}\frac{n_i}{r^2}$$

we finally get

$$\oint \sqrt{|g|}\, g^{\rho A}\Gamma^i_{A\sigma} n_i\, df = 2\pi\epsilon_\rho m_{\rho\sigma}\,. \tag{22}$$

According to this equation, the integral theorems (13a) imply

$$m_{\rho\sigma} = 0 \quad \text{for all} \quad \rho, \sigma, \tag{23}$$

for every solution which is independent of the x^ρ, regular everywhere, and asymptotically approaches the Euclidean metric. Moreover, the relation (20) leads to

$$m = 0. \tag{23a}$$

This shows that for such a non-singular solution of the field equations $R_{IK} = 0$ the deviations of the g_{IK} from the Euclidean (or Minkowskian) $\mathring{g}_{IK}$ must decrease more rapidly than $1/r$ for all values of I and K. Should there exist at all a solution different from the Euclidean metric, it could not describe a particle with *non-vanishing mass or charge*, as stated in the introduction.

Appendix

In the special case of the four-dimensional continuum, R. Serini [Atti Academia dei Lincei (5), 27[1], 235, 1918] has shown that except for the Euclidean metric there exists no regular solution of the required form, under the restrictive assumption, however, that the g_{i4} vanish everywhere. His proof is based on the fact that, under this assumption, we get for the equation (13)

$$(\sqrt{|g|}\, g^{44} g^{ik} g_{44,k})_{,i} = 0.$$

Multiplying this by $-g_{44}$ and integrating over a three-dimensional region we find, because of $g^{44}g_{44} = -1$,

$$\int \sqrt{|g|}\, g^{44} g^{ik} g_{44,i} g_{44,k}\, d^3x + \oint \sqrt{|g|}\, g^{ik} g_{44,k} n_i\, df = 0.$$

According to (13a), the surface integral approaches 0 as the surface moves to infinity, because then $g_{44} \to -1$, $g_{ik} \to \delta_{ik}$. Since $g^{44} \neq 0$, $g \neq 0$, and the form $g^{ik}x_i x_k$ is positive definite, the vanishing of the volume integral (extended over the whole space) implies $g_{44} = \text{const}$. For a three-dimensional space, however, the field equations $R_{ik} = 0$ are equivalent to the requirement that the uncontracted Riemannian curvature tensor vanish; hence they imply that the space is Euclidean.

It seems impossible to treat those stationary solutions of the field equations for which the γ_{IK} [i.e. the deviations of the g_{IK} from the constant values $\mathring{g}_{IK}$] decrease more rapidly than $1/r$ by the methods applied in this paper, namely by the use of integral theorems and the linearized field equations. In order to investigate these solutions it will be necessary to discuss in more detail the higher approximations or the exact field equations.

INSTITUTE FOR ADVANCED STUDY

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April 1943

ON THE SUM OF TWO SETS OF INTEGERS

BY EMIL ARTIN AND PETER SCHERK

(Received January 30, 1943)

In his beautiful paper: "A proof of the fundamental theorem on the density of sums of sets of positive integers"[1] Mr. Mann succeeded in proving the (α, β)-hypothesis and a generalization of it that had been conjectured for more than ten years. We found that his method can be simplified considerably and even yields some stronger results.

Let A, B respectively be sets of nonnegative integers a, b. Let $C = A + B$ be the set of all integers of the form $a + b$. Let $A(x)$, $B(x)$, $C(x)$ denote the number of positive integers of the sets $\leqq x$.[2] Mr. Mann proved the following theorem:

If $0 \subset A$ *and* $0 \subset B$, *and if* $C(n) < n$, *then*

$$\frac{C(n)}{n} \geqq \operatorname*{Min}_{\substack{x=1,\cdots,n \\ x \text{ not in } C}} \frac{A(x) + B(x)}{x}. \tag{1}$$

It seems to us remarkable that it is possible to prove a certain identity from which Mr. Mann's inequality can be deduced as an immediate consequence. The theorem in question is the following:

THEOREM I. *Let* $n \notin C$. *Then*

$$C(n) - C(n - m) = A(m - 1) + B(m - 1) + Z_m \tag{2}$$

for a suitable $m \notin C$, $0 < m \leqq n$, *where* Z_m *denotes the number of decompositions of* m *of a certain type.*

Throughout the following proof, small letters always denote numbers between 0 and n, and capital letters stand for sets of such numbers; n is supposed to be not in C. We construct now several sequences of sets denoted by B_ν, B_ν^*, C_ν, and C_ν^*.

Let $C_0 = C_0^* = C$, $B_0 = B_0^* = B$. Let e_1 be the smallest number of B_0^* for which there are two numbers c_1, c_1' not in C_0^* such that

$$c_1 + c_1' - n = a + e_1. \tag{3}$$

With this e_1, we now form C_1 as the set of solutions c_1, c_1' of (3). The corresponding numbers

$$e_1 + n - c_1 = c_1' - a,$$

form the set B_1. Such a number e_1 need not exist; in this case, our construction stops at C_0. C_1 exists if and only if there are elements in C_0^* that have the form $f + f' - n$, where $f, f' \notin C_0^*$. The sets C_1^* and B_1^* shall be the union of C_0^* and C_1 respectively of B_0^* and B_1.

[1] Annals of Math. 43(1942), pp. 523–527.

[2] Thus, 0 is never counted.

LEMMA 1. $A + B_1^* = C_1^*$.[3]

PROOF: The elements of B_1 can be written in the form $c_1' - a$. So, $A + B_1 \supset C_1$. Since $A + B_0 = C = C_0$, we have $A + B_1^* \supset C_1^*$. So, we have to show that every number $a + b_1 = a + (e_1 + n - c_1)$ belongs to C_1^*, provided it is less than n. In proving this, we may assume that this number does not lie in C_0^*. Calling it c_1', it turns out to be a solution of (3) and, hence, to belong to C_1.

LEMMA 2. $n \notin C_1^*$.

It is sufficient to prove $n \notin C_1$. Suppose $n = c_1$ in (3) then $c_1' = a + e_1$; from $e_1 \subset B_0^*$ would follow $c_1' \subset C_0^*$.

LEMMA 3. *B_1 and B_0^* are disjoint.*

For if $e_1 + n - c_1 = c_1' - a \subset B_0^*$, c_1' would be of the form $a + b \subset C_0^*$.

Starting from B_1^* and C_1^*, we construct in the same way B_2^*, C_2^*, B_3^*, C_3^*, and so on. This is possible, because B_1^*, C_1^* satisfy still the same conditions as B_0^* and C_0^*, namely Lemma 1 and 2. This process stops, say, at C_h^*. Thus, no number of C_h^* has the form $f + f' - n$, where $f, f' \notin C_h^*$. The corresponding numbers e shall be called $e_2, e_3, \cdots, e_h$. Thus, $e_\nu \subset B_{\nu-1}^*$; $n \notin C_h^*$, and any two sets B_ν are disjoint.

LEMMA 4. *The numbers e_ν increase monotonically.*[4]

It suffices to show that $e_2 > e_1$. If $e_2 \subset B_0^*$, this follows from the minimum definition of e_1. If $e_2 \subset B_1$, then $e_2 = e_1 + (n - c_1) > e_1$, on account of Lemma 2.

We define $m \leqq n$ as the smallest positive number not in C_h^*.

LEMMA 5. *There are no numbers $c_1 \subset C_1$ with $n - m < c_1 \leqq n - m + e_1$.*

PROOF. Let $c_1 \subset C_1$, $n - m < c_1$. We wish to show $n - m + e_1 < c_1$. This contention is equivalent to $m + c_1 - n > e_1$. Obviously, $0 < m + c_1 - n < m$, hence $m + c_1 - n \subset C_h^*$, say $m + c_1 - n \subset C_\nu$; therefore, because of $C_\nu \subset A + B_\nu$: $m + c_1 - n = a + b_\nu \geqq b_\nu$.

If $\nu > 0$, then $b_\nu = e_\nu + (n - c_\nu'') > e_\nu \geqq e_1$. If $\nu = 0$, we obtain $m + c_1 - n = a + b$; $c_1' = m$ and c_1 are a solution of (3) with b instead of e_1. Since e_1 was chosen minimal, and since $m \notin C_1$, we obtain $b > e_1$.

There are $C_1(n) - C_1(n - m)$ numbers c_1 in the interval $n - m < c_1 < n$. According to Lemma 5, they even satisfy $n - m + e_1 < c_1$ or $e_1 + n - c_1 < m$. These are, according to the definition of B_1, precisely the numbers of B_1 below m. Their number is equal to $B_1(m - 1)$. Thus, we obtain for any $\nu \geqq 1$

$$C_\nu(n) - C_\nu(n - m) = B_\nu(m - 1). \tag{4}$$

LEMMA 6. *All the numbers s in the interval $n - m < s < n$ belong to C_h^*.*

They satisfy, indeed, $0 < s + m - n < m$, so that $s + m - n \subset C_h^*$. If s would not belong to C_h^*, we could construct C_{h+1}^*, for m is also not in C_h^*.

Since $n \notin C_h^*$, we obtain from Lemma 6

$$C_h^*(n) - C_h^*(n - m) = m - 1. \tag{5}$$

[3] The sum-sign means: C_1^* is the set of all numbers $a + b_1^*$, where $a \subset A$ and $b_1^* \subset B_1$.

[4] We owe this lemma to a written communication of Professor Alfred Brauer.

There are $m-1$ decompositions of m into positive summands: $m = x + y$. Among them, there are $A(m-1)$ with $x \subset A$ and $B_h^*(m-1)$ with $y \subset B_h^*$. Since $m \notin A + B_h^*$, these $A(m-1) + B_h^*(m-1)$ decompositions are different from each other. If Z_m denotes the number of decompositions $m = x + y$, where $x \notin A$ and $y \notin B_h^*$, then

$$m - 1 = A(m-1) + B_h^*(m-1) + Z_m .$$

According to Lemma 3, the sets B_ν are disjoint, hence

$$B_h^*(m-1) = \sum_0^h B_\nu(m-1)$$

and

$$m - 1 = A(m-1) + \sum_0^h B_\nu(m-1) + Z_m . \tag{6}$$

On the other hand, the C_ν are disjoint by construction. So, (5) may be written in the form

$$m - 1 = \sum_0^h (C_\nu(n) - C_\nu(n-m)),$$

or using (4) for $\nu = 1, \cdots, h$

$$m - 1 = C(n) - C(n-m) + \sum_1^h B_\nu(m-1). \tag{7}$$

By comparing the right sides of (6) and (7), we obtain (2).

As a consequence of Theorem I, we obtain

THEOREM II. *If $C(n) < n$, then*

$$C(n) - C(n-m) \geqq A(m-1) + B(m-1) \tag{8}$$

for a suitable $m \notin C$ with $0 < m \leqq n$.

Obviously, (2) implies (8) for $n \notin C$. If (8) holds for $n-1$, and if $n \subset C$, then we choose the same m for n as for $n-1$; the right term of (8) remains unchanged, when we replace $n-1$ by n, while the left term is not decreased.

If $0 \subset A$, $0 \subset B$, then $m \notin B$ and $m \notin A$; for if $m \subset A$, we would have $m = 0 + m \subset C$. So, in this case, (8) implies

$$C(n) - C(n-m) \geqq A(m) + B(m). \tag{9}$$

Iterating this formula, we can obtain the following inequality

$$C(n) \geqq \operatorname*{Min}_{\substack{n_0 + \sum m_i = n \\ m_i \text{ not in } C}} (n_0 + \sum (A(m_i) + B(m_i))). \tag{10}$$

Mr. Mann's estimate (1) is an immediate consequence of either (9) or (10).

For

$$n_0 \geqq n_0 \operatorname*{Min}_{\substack{x=1,\cdots,n \\ x \text{ not in } C}} \left(1, \frac{A(x)+B(x)}{x}\right),$$

$$A(m_i)+B(m_i) = m_i \frac{A(m_i)+B(m_i)}{m_i} \geqq m_i \operatorname*{Min}_{\substack{x=1,\cdots,n \\ x \text{ not in } C}} \left(1, \frac{A(x)+B(x)}{x}\right);$$

hence, on account of

$$n_0 + \sum m_i = n,$$

$$C(n) \geqq n \cdot \operatorname*{Min}_{\substack{x=1,\cdots,n \\ x \text{ not in } C}} \left(1, \frac{A(x)+B(x)}{x}\right). \tag{11}$$

Since $C(n) < n$ by assumption,

$$\operatorname*{Min}_{\substack{x=1,\cdots,n \\ x \text{ not in } C}} \left(1, \frac{A(x)+B(x)}{x}\right) = \operatorname*{Min}_{\substack{x=1,\cdots,n \\ x \text{ not in } C}} \frac{A(x)+B(x)}{x},$$

and (1) follows from (11).

Another consequence of (10) is the (α, β)-theorem: Let

$$\alpha > 0, \quad \beta > 0, \quad \alpha + \beta < 1; \quad A(x) \geqq \alpha x, \quad B(x) \geqq \beta x \quad \text{for } x = 1, 2, \cdots, n. \tag{12}$$

Then

$$C(n) \geqq (\alpha+\beta)n.$$

This is clear if $C(n) = n$. But if $C(n) < n$, then (10) yields, on account of $n_0 \geqq (\alpha+\beta)n_0$,

$$C(n) \geqq \operatorname*{Min}_{\substack{n_0+\sum m_i = n \\ m_i \text{ not in } C}} ((\alpha+\beta)n_0 + \sum (\alpha m_i + \beta m_i))$$

$$= \operatorname*{Min}_{\substack{n_0+\sum m_i = n \\ m_i \text{ not in } C}} (\alpha+\beta)(n_0 + \sum m_i) = (\alpha+\beta)n.$$

Obviously, (12) can be replaced by the weaker assumption

$$0 < \gamma < 1, \quad A(x) + B(x) \geqq \gamma x \quad \text{for} \quad x = 1, \cdots, n.$$

Let $0 \subset A$, $0 \not\subset B$. Then C consists of all numbers of the form b and $b + a$, where $a \subset A$, $b \subset B$, and both positive. For such sets, Mr. A. S. Besicovitch has proved[5]: If $A(x) \geqq \alpha(x+1)$, $B(x) \geqq \beta x$ for $x = 1, \cdots, n$, and if $C(n) < n$, then $C(n) \geqq (\alpha+\beta)n$.

[5] *A. S. Besicovitch*: *On the density of the sum of two sets of integers*, Journal London Math. Soc., vol. 10 (1935), pp. 246–248. Mr. Besicovitch's method yields the stronger result: If $A(x) \geqq \alpha(x+1)$ for $x = 1, \cdots, n$ and if $C(n) < n$, then $C(n) \geqq B(n) + \alpha n$.

Since, in this case, (8) goes over into

$$C(n) - C(n - m) \geqq A(m - 1) + B(m),$$

Mr. Besicovitch's estimate is a consequence of Theorem II. In this case, Theorem II implies inequalities that are analogous to those discussed above.

INDIANA UNIVERSITY

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

CONTRIBUTIONS TO THE THEORY OF THE DIRICHLET L-SERIES AND THE EPSTEIN ZETA-FUNCTIONS

BY CARL LUDWIG SIEGEL

Introduction

Let

$$t > 0, \qquad \nu = \left[\sqrt{\frac{t}{2\pi}}\right],$$

$$\vartheta = -\frac{t}{2}\log\pi + \arg\Gamma\left(\frac{1}{4} + \frac{ti}{2}\right) = \frac{t}{2}\log\frac{t}{2\pi e} - \frac{\pi}{8} + O(t^{-1}),$$

$$e^{i\vartheta}\zeta(\tfrac{1}{2} + ti) - 2\sum_{n=1}^{\nu} n^{-\frac{1}{2}}\cos(\vartheta - t\log n) = R.$$

The formula

$$R = O(t^{-\frac{1}{4}}), \tag{1}$$

due to Hardy and Littlewood[1], is important in the theory of the zeta-function. This formula contains the main term of an asymptotic expansion of $\zeta(\frac{1}{2} + ti)$ for $t \to \infty$, which had been discovered already by Riemann, but was not published[2] before 1932. Riemann's formula is

$$R = (-1)^{\nu-1}\left(\frac{2\pi}{t}\right)^{\frac{1}{4}}\left(C_0 + C_1 t^{-\frac{1}{2}} + \cdots + C_k t^{-\frac{k}{2}} + R_k\right), \qquad R_k = O\left(t^{-\frac{k+1}{2}}\right), \tag{2}$$

where k is an arbitrary integer $\geqq 0$ and $C_0, C_1, \cdots, C_k$ denote certain bounded functions of t, e.g. $C_0 = \cos\left(2\pi u^2 + \frac{3\pi}{8}\right)\Big/\cos 2\pi u$, $u = \sqrt{\frac{t}{2\pi}} - \nu - \frac{1}{2}$. This has been used by Titchmarsh[3], with $k = 1$ and numerical bounds of R_1, for the calculation of the zeros of $\zeta(\sigma + ti)$ in the strip $0 < t < 1468$; he found that all 1041 zeros lie on the critical line $\sigma = \frac{1}{2}$.

Kusmin[4] generalized (1) for the case of an arbitrary L-series, $L(s) = \sum_{n=1}^{\infty}\chi(n)n^{-s}$, where $\chi(n)$ denotes a proper character modulo $m \geqq 1$. In the

[1] G. H. Hardy and J. E. Littlewood, *The zeros of Riemann's zeta-function on the critical line*, Math. Zeitschr. 10, pp. 283–317 (1921).

[2] C. L. Siegel, *Über Riemanns Nachlass zur analytischen Zahlentheorie*, Quell. u. Stud. z. Gesch. d. Math. B2, pp. 45–80 (1932).

[3] E. C. Titchmarsh, *The zeros of the Riemann zeta-function*, Proc. Roy. Soc. London A 151, pp. 234–255 (1935), and 157, pp. 261–263 (1936).

[4] R. O. Kusmin, *Sur les zéros de la fonction $\zeta(s)$ de Riemann*, C. R. Acad. Sci. URSS (N.S.) 2, pp. 398–400 (1934) (Russian. French summary.); R. O. Kusmin, *Zur Theorie der Dirichletschen Reihen $L(s)$*, Bull. Acad. Sci. URSS (7), pp. 1471–1491 (1934) (Russian. German summary).

present paper, I prove the analogue of (2) for $L(s)$; the result (Theorem 6) is of the same form as (2) and contains, of course, (2) as the special case $m = 1$. My proof is somewhat simpler than my former proof of (2). It starts from a representation of $L(s)$ as the sum of two particular integrals (Theorem 4), obtained in a different way by Kusmin; the corresponding theorem for $\zeta(s)$, discovered by Riemann, was published in my edition of Riemann's manuscripts on analytical number-theory. Using a simplification of my former method, remarked by Kusmin, I prove then the following two theorems:

Let $A(t_1, t_2)$ denote the number of different zeros of odd order of $L(s)$ in the interval $t_1 < t < t_2$ on the critical line $s = \frac{1}{2} + ti$. If t_1 is a function of t satisfying the condition $t^{\frac{1}{3}} \log t = o(t - t_1)$, then

$$\liminf_{t\to\infty} A(t_1, t)/(t - t_1) \geqq \gamma, \tag{3}$$

where $\gamma = m/\pi e\varphi(m)$ and $\varphi(m)$ is Euler's function.

Let $B(t_1, t_2, \epsilon)$ denote the number of zeros of $L(s)$ in the rectangle $t_1 < t < t_2$, $\frac{1}{2} - \epsilon < \sigma < \frac{1}{2} + \epsilon$. If $t^{\frac{1}{3}} \log t = o(t - t_1)$ and $\epsilon = o(\log\log t/\log t)$, then

$$\liminf_{t\to\infty} B(t_1, t, \epsilon)/(t - t_1)t^{\epsilon} \geqq \tfrac{1}{4}\gamma. \tag{4}$$

Hardy and Littlewood proved that (3) holds in the case of $\zeta(s)$ with a positive constant γ, provided $t^{\lambda} \ll O(t - t_1)$ with constant $\lambda > \frac{1}{2}$; however, they did not determine an explicit value of γ. The value $\gamma = m/\pi e\varphi(m)$ is better than the values formerly obtained, in the case $t_1 = 0$, by me for $\zeta(s)$ and by Kusmin for an arbitrary L-function.

The interest of (4) consists in the condition $\epsilon = o(\log\log t/\log t)$ for the breadth of the rectangle. In the well-known theorems of Littlewood[5] and Carlson[6], this breadth is subjected to the conditions $\log\log t/\log t = o(\epsilon)$ and $1 = O(\epsilon)$.

The second part of the paper is concerned with similar problems for certain Epstein zeta-functions, whereas the methods are quite different. Let $Q(x) = Q(x_1, \cdots, x_k)$ be a positive quadratic form of k variables, $\mathfrak{S}$ its matrix, D its determinant and $\zeta(s; \mathfrak{S})$ the Epstein zeta-function defined by the series $\zeta(s; \mathfrak{S}) = \sum Q(n)^{-s}$ $\left(\sigma > \frac{k}{2}\right)$, where $n = (n_1, \cdots, n_k)$ runs over all lattice-points in the k-dimensional space with exception of the origin. The function $\left(s - \frac{k}{2}\right)\zeta(s; \mathfrak{S})$ is regular in the whole plane, and $\eta(s; \mathfrak{S}) = \pi^{-s}\Gamma(s)\zeta(s; \mathfrak{S})$ fulfills the functional equation $\eta(s; \mathfrak{S}) = D^{-\frac{1}{2}}\eta\left(\frac{k}{2} - s; \mathfrak{S}^{-1}\right)$. Obviously $\zeta(s; \mathfrak{C}'\mathfrak{S}\mathfrak{C}) = \zeta(s; \mathfrak{S})$, for any unimodular matrix $\mathfrak{C}$ with k rows; hence $\zeta(s; \mathfrak{S}) = \zeta(s; \mathfrak{S}^{-1})$, if $\mathfrak{S}$ itself

[5] J. E. Littlewood, *On the zeros of the Riemann zeta-function*, Proc. Cambridge Philos. Soc. 22, pp. 295–318 (1925).

[6] F. Carlson, *Über die Nullstellen der Dirichlet'schen Reihen und der Riemann'schen ζ-Funktion*, Ark. Mat. Astr. Fys. no. 20, 28 pp. (1921).

is unimodular. The functional equation shows, in this case, that $\eta(s; \mathfrak{S})$ is real on the line $\sigma = \frac{k}{4}$, which corresponds to the critical line $\sigma = \frac{1}{2}$ for the zeta-function. This holds in particular for $\mathfrak{S} = \mathfrak{E}_k$, the unit matrix of k rows; put $\zeta(s; \mathfrak{E}_k) = \zeta_k(s)$. The formula $\zeta_1(s) = 2\zeta(2s)$ suggests that the distribution of the zeros of $\zeta_k(s)$ might be analogous to that of $\zeta(s)$ itself. Generalizing Hardy's first proof of the existence of an infinite number of zeros of $\zeta(\frac{1}{2} + ti)$, Landau[7] obtained the same result for all $\zeta_k\left(\frac{k}{4} + ti\right)$. He did not notice that, in the special cases $k = 4, 8$, his theorem is an immediate consequence of the formulae

(5) $\zeta_4(s) = 8(1 - 2^{2-2s})\zeta(s)\zeta(s-1)$, $\zeta_8(s) = 16(1 - 2^{1-s} + 2^{4-2s})\zeta(s)\zeta(s-3)$

following from Jacobi's theorems on the number of decompositions of an integer into 4 and 8 squares; obviously $\zeta_4(s)$, $\zeta_8(s)$ have on the critical line $\sigma = \frac{k}{4}$ exactly the zeros $s = 1 \pm \frac{l\pi i}{\log 2}$ $(l = 1, 2, \cdots)$, $s = 2 + \frac{i}{\log 2}(2l\pi \pm \operatorname{arc\,tg}\sqrt{15})$ $(l = 0, \pm 1, \cdots)$. For other values of k, however, there is no such obvious reason for the existence of the zeros. My results depend upon two theorems concerning the Epstein zeta-function for arbitrary integral $\mathfrak{S}$:

If $\mathfrak{S}$ and $\mathfrak{S}_1$ belong to the same genus, then

$$\zeta(s; \mathfrak{S}) - \zeta(s; \mathfrak{S}_1) = t^{\frac{1}{2}} \log t\, O\left(1 + t^{\frac{k}{2}-2\sigma}\right). \tag{6}$$

Let $\mathfrak{S}_1, \cdots, \mathfrak{S}_h$ be representatives of the different classes of the genus of $\mathfrak{S}$, let $E(\mathfrak{S})$ be the number of units of $\mathfrak{S}$ and

$$Z(s) = \sum_{l=1}^{h} \frac{\zeta(s; \mathfrak{S}_l)}{E(\mathfrak{S}_l)} \Big/ \sum_{l=1}^{h} \frac{1}{E(\mathfrak{S}_l)};$$

define

$$H\left(\frac{a}{b}\right) = D^{-\frac{1}{2}} b^{-\frac{k}{2}} \sum_{n \,(\mathrm{mod}\, b)} e^{\pi i \frac{a}{b} Q(n)},$$

where a, b are coprime positive integers and $abQ(x)$ is an even quadratic form; then

$$Z(s) = \pi^s \frac{\Gamma\left(\frac{k}{2} - s\right)}{\Gamma\left(\frac{k}{2}\right)} \sum_{a,b} a^{s-\frac{k}{2}} b^{-s} \left\{ e^{\frac{\pi i}{4}(2s-k)} H\left(\frac{a}{b}\right) + e^{\frac{\pi i}{4}(k-2s)} \overline{H\left(\frac{a}{b}\right)} \right\} \tag{7}$$

$$\left(1 < \sigma < \frac{k}{2} - 1\right).$$

[7] E. Landau, *Über die Hardysche Entdeckung unendlich vieler Nullstellen der Zetafunktion mit reellem Teil* $\frac{1}{2}$, Math. Ann. 76, pp. 212–243 (1915).

The proofs of these two theorems use the theory of modular forms and the analytic theory of quadratic forms. The formulae (6), (7) contain an analogue of (1); it is easy to deduce the following statements:

Let $A(t_1)$ denote the number of different zeros of odd order of $\zeta(s; \mathfrak{S})$ in the interval $0 < t < t_1$ on the line $s = \frac{k}{4} + ti$. If $\mathfrak{S}$ belongs to the genus of $\mathfrak{E}_k$, then

$$A(t) > \frac{t}{\pi} \log 2 + O(1) \qquad (k > 8). \tag{8}$$

Let $B(t_1)$ denote the number of zeros of $\zeta(s; \mathfrak{S})$ in the rectangle $0 < t < t_1$, $2 \leqq \sigma \leqq \frac{k}{2} - 2$. If $\mathfrak{S}$ belongs to the genus of $\mathfrak{E}_k$, then

$$B(t) = \frac{t}{\pi} \log 2 + O(1) \qquad (k \geqq 12). \tag{9}$$

Obviously, (8) and (9) correspond to (3) and (4). The consequences are much more precise; it follows immediately, for $k \geqq 12$, that the zeros of $\zeta(s; \mathfrak{S})$ in the strip $2 \leqq \sigma \leqq \frac{k}{2} - 2$ are simple and lie on $\sigma = \frac{k}{4}$, with at most a finite number of exceptions. In the cases $k = 4, 8$, the zeros of $\zeta(s; \mathfrak{S})$ on $\sigma = \frac{k}{4}$ are completely known, by (5); it is possible to discuss also the remaining cases for $3 < k < 12$, but this requires some numerical computations and we omit it. For the function $\zeta_3(s)$, however, our method does not lead to any result.

Since the number $N(t_1)$ of all zeros in the strip $0 < t < t_1$ satisfies $N(t) = \frac{t}{\pi} \log t + O(t)$, it follows that most of the zeros of $\zeta(s; \mathfrak{S})$ do not lie in the neighborhood of $\sigma = \frac{k}{4}$, if $\mathfrak{S}$ belongs to the genus of $\mathfrak{E}_k$ and $k \geqq 12$.

Part I: Dirichlet L-series

1. Asymptotic expansion

Let σ_1, σ_2 be given real numbers, $\sigma_1 < \sigma_2$, and $s = \sigma + ti$ a complex variable in the half-strip $\sigma_1 \leqq \sigma \leqq \sigma_2$, $t \geqq 2$. If $P, Q \neq 0$ are functions of s and some parameters $p, u, n, \cdots$, then the formula $P = O(Q)$ means that P/Q is bounded in the half-strip, uniformly with respect to the set of values of all parameters $p, u, \cdots$ except n. The symbols $\Re\{c\}$, $\Im\{c\}$ denote real and imaginary part of a complex number c, and $\bar{c}$ is the conjugate complex number. We define

$$\tau = \sqrt{i\left(\frac{1}{2} - s\right)} = \sqrt{t + i\left(\frac{1}{2} - \sigma\right)} = \sqrt{t}\,(1 + O(t^{-1})),$$

$$\sqrt{t} > 0, \qquad \epsilon = e^{\frac{\pi i}{4}} = \frac{1+i}{\sqrt{2}}, \qquad g(z) = e^{\frac{i}{2}z^2} z^{-s}$$

with the principal value of $z^{-s} = e^{-s \log z}$; moreover u is a real variable, p is a positive parameter satisfying the condition $p = \tau + O(1)$ and $c_1, \cdots, c_{21}$ are certain appropriate positive constants.

LEMMA 1:

$$g(p + \epsilon u)/g(\tau) = O(e^{-c_1 u^2}).$$

PROOF: Let $\tau^{-1}(\epsilon u + p - \tau) = v = |v| e^{i\alpha}$, $-\pi \leqq \alpha < \pi$; then

$$v = t^{-\frac{1}{2}}(\epsilon u + O(1))(1 + O(t^{-1})), \tag{10}$$

whence $\alpha = \arg v = -\frac{\pi}{4} + \frac{\pi}{2} \operatorname{sign} u + O(u^{-1}) + O(t^{-1})$, $\cos \alpha > -\frac{3}{4}$ ($u^2 > c_2$; $t > c_3$) and $(\tau v)^2 = iu^2 + (|u| + 1)\, O(1)$; moreover

$$|1 + v| = \frac{p}{|\tau|}\left|1 + \frac{\epsilon}{p} u\right| \geqq \frac{p}{|\tau| \sqrt{2}} > c_4 \qquad (t > c_5). \tag{11}$$

Put

$$\lambda = v + \tfrac{1}{2}v^2 - \log(1 + v) = \int_0^v \frac{w + 2}{w + 1} w\, dw, \qquad e^{-2i\alpha}\lambda = \mu,$$

where the integration is performed over the segment $w = re^{i\alpha}$, $0 \leqq r \leqq |v|$. By (11), $\mu = v^2 O(1)$; on the other hand $\left(r^2 - \frac{9}{4} r + 2\right)(1 + r)^{-2} > c_6$ $(0 \leqq r)$,

$$\begin{aligned} \Re\{\mu\} &= \int_0^{|v|} \frac{r^2 + 3r \cos \alpha + 2}{r^2 + 2r \cos \alpha + 1} r\, dr \\ &\geqq \int_0^{|v|} \left(r^2 - \frac{9}{4} r + 2\right)(1 + r)^{-2} r\, dr \geqq \tfrac{1}{2} c_6 |v|^2 \qquad (u^2 > c_2;\ t > c_3). \end{aligned}$$

Therefore

$$\begin{aligned} \Im\{\tau^2 \lambda\} &= \Im\{\tau^2 v^2 |v|^{-2} \mu\} = u^2 |v|^{-2} \Re\{\mu\} + (|u| + 1)|v|^{-2} \mu\, O(1) \\ &\geqq \tfrac{1}{2} c_6 u^2 + O(|u| + 1) > c_1 u^2 \qquad (u^2 > c_7;\ t > c_3) \end{aligned} \tag{12}$$

and

$$\Im\{\tau^2 \lambda\} = O(1) \qquad (u^2 \leqq c_7). \tag{13}$$

Since $g(p + \epsilon u) = g(\tau)(1 + v)^{-\frac{1}{2}} e^{i\tau^2 \lambda}$, the assertion follows from (11), (12), (13).

Consider now v as an independent complex variable and define

$$h = h(v) = v - \tfrac{1}{2}v^2 - \log(1 + v) \tag{14}$$

with the principal value of $\log(1 + v)$,

$$\psi(v) = (1 + v)^{-\frac{1}{2}} e^{i\tau^2 h}. \tag{15}$$

The function $\psi(v)$ is regular in the circle $|v| < 1$, let

$$\psi(v) = \sum_{n=0}^{\infty} A_n v^n \qquad (|v| < 1)$$

be its power series and

$$S_n(v) = \sum_{k=0}^{n-1} A_k v^k, \qquad R_n(v) = \psi(v) - S_n(v) \qquad (n = 0, 1, \cdots).$$

LEMMA 2: *Let u be real and $v = \tau^{-1}(\epsilon u + p - \tau)$; then*

$$R_n(v) = (|u| + 1)^n O(t^{-\frac{n}{6}}) \quad (u^2 \leqq c_8 t^{\frac{1}{3}}), \qquad S_n(v) = O(u^n) \quad (u^2 \geqq c_8 t^{\frac{1}{3}}).$$

PROOF: By (14), the function $v^{-3}h(v)$ is regular in the circle $|v| < 1$ and consequently bounded for $|v| \leqq \frac{1}{2}$, hence $\tau^2 h(v) = O(1)$, $\psi(v) = O(1)$ $(|v| \leqq \frac{1}{2}t^{-\frac{1}{3}})$. Applying the formula

$$|R_n(v)| = \left| \frac{1}{2\pi i} \int_{|z|=\rho} \frac{v^n \psi(z)}{z^n (z - v)}\, dz \right| \leqq \left(\frac{|v|}{\rho}\right)^n \frac{\rho}{\rho - |v|} \max_{|z|=\rho} |\psi(z)| \quad (|v| < \rho < 1)$$

with $\rho = \frac{1}{2}t^{-\frac{1}{3}}$, $|v| < \frac{1}{2}\rho$, we obtain

$$R_n(v) = v^n O\left(t^{\frac{n}{3}}\right) \qquad (|v| < c_9 t^{-\frac{1}{3}}). \tag{16}$$

Moreover $A_n = v^{-n}(R_{n+1} - R_n) = O\left(|v|\, t^{\frac{n+1}{3}} + t^{\frac{n}{3}}\right)$ $(|v| < c_9 t^{-\frac{1}{3}})$, whence $A_n = O\left(t^{\frac{n}{3}}\right)$,

$$S_n(v) = O\left(1 + |v|^n t^{\frac{n}{3}}\right). \tag{17}$$

By (10), the condition $|v| < c_9 t^{-\frac{1}{3}}$ is satisfied for $v = \tau^{-1}(\epsilon u + p - \tau)$, whenever $u^2 \leqq c_8 t^{\frac{1}{3}}$. The assertion follows from (10), (16), (17).

LEMMA 3: *The coefficients $A_n = A_n(\tau)(n = 0, 1, \cdots)$ of the power series $\psi(v)$ are polynomials in τ^2 of degree $\leqq \frac{n}{3}$ satisfying the recursion formula $nA_n + (n - \frac{1}{2})A_{n-1} + i\tau^2 A_{n-3} = 0$ $(n = 1, 2, \cdots)$ with $A_0 = 1$, $A_{-1} = A_{-2} = 0$.*

PROOF: The function $\psi = \psi(v)$ fulfills the differential equation

$$\frac{d \log \psi}{dv} = -\frac{1}{2}\frac{1}{1 + v} + i\tau^2\left(1 - v - \frac{1}{1 + v}\right),$$

whence

$$(1 + v)\frac{d\psi}{dv} + (\tfrac{1}{2} + i\tau^2 v^2)\psi = 0,$$

$nA_n + (n - 1)A_{n-1} + \frac{1}{2}A_{n-1} + i\tau^2 A_{n-3} = 0$ $(n = 1, 2, \cdots)$, moreover $A_0 = \psi(0) = 1$; q.e.d.

Let $f(u)$ be an integrable function of the real variable u and

$$|f(u)| < c_{10} e^{c_{11} u^2}, \qquad c_{11} < c_1, c_{11} < 1. \tag{18}$$

Setting $\epsilon u + p = z$, we define

$$F(s) = \int_{-\infty}^{+\infty} g(z) f(u)\, du, \qquad B_n = \int_{-\infty}^{+\infty} e^{i(z-\tau)^2}(z - \tau)^n f(u)\, du \qquad (n = 0, 1, \cdots).$$

LEMMA 4:

$$F(s)/g(\tau) = \sum_{k=0}^{n-1} A_k B_k \tau^{-k} + O\left(t^{-\frac{n}{6}}\right) \qquad (n = 0, 1, \cdots).$$

PROOF: Introducing again $v = \tau^{-1}(\epsilon u + p - \tau) = \tau^{-1}(z - \tau)$, we have, by (14), (15), $g(z)/g(\tau) = e^{i(z-\tau)^2}\psi(v)$. We use the decomposition

$$F(s)/g(\tau) = \int_{-\infty}^{+\infty} e^{i(z-\tau)^2} S_n(v) f(u)\, du - \int_{|u| \geqq a} e^{i(z-\tau)^2} S_n(v) f(u)\, du$$

$$+ \int_{-a}^{a} e^{i(z-\tau)^2} R_n(v) f(u)\, du + \int_{|u| \geqq a} g(z) f(u)\, du/g(\tau)$$

with $a^2 = c_8 t^{\frac{1}{3}}$ and obtain, by (18) and Lemmata 1, 2,

$$F(s)/g(\tau) - \sum_{k=0}^{n-1} A_k B_k \tau^{-k} = \int_a^{\infty} e^{-c_{12}u^2} O(u^n)\, du$$

$$+ \int_0^a e^{-c_{12}u^2}(u+1)^n O\left(t^{-\frac{n}{6}}\right) du + \int_a^{\infty} O(e^{-c_{12}u^2})\, du$$

$$= O(e^{-c_{14}t^{\frac{1}{3}}}) + O\left(t^{-\frac{n}{6}}\right) + O(e^{-c_{15}t^{\frac{1}{3}}}) = O\left(t^{-\frac{n}{6}}\right);$$

q.e.d.

We write $P \approx \sum_{n=0}^{\infty} Q_n \tau^{-n}$ as an abbreviation for the formula $P - \sum_{k=0}^{n-1} Q_k \tau^{-k} = O(\tau^{-n})(n = 0, 1, \cdots)$. It is easily seen that the relation $P \approx Q$ has the following simple properties: If $P \approx Q$ and $P^* \approx Q^*$, then $P + P^* \approx Q + Q^*$ and $PP^* \approx QQ^*$, where $Q + Q^*$ and QQ^* are sum and product of the formal power series Q, Q^*; if $P \approx Q$, then $P^{-1} \approx Q^{-1}$, whenever the first coefficient Q_0 of Q satisfies the condition $Q_0^{-1} = O(1)$.

THEOREM 1:

$$F(s)/g(\tau) \approx \sum_{k=0}^{\infty} \Gamma_k \tau^{-k},$$

where

$$\Gamma_k = \sum_{l=0}^{3k} \gamma_{kl} \Phi^{(l)}(\tau), \qquad \Phi(y) = \int_{-\infty}^{+\infty} e^{i(z-y)^2} f(u)\, du$$

with numerical coefficients γ_{kl}; *in particular* $\gamma_{00} = 1$.

PROOF: Since

$$e^{-iy^2}\Phi(\tau + y) = \int_{-\infty}^{+\infty} e^{i(z-\tau)^2 - 2iy(z-\tau)} f(u)\, du,$$

we find

$$(19) \qquad B_k = \int_{-\infty}^{+\infty} e^{i(z-\tau)^2}(z-\tau)^k f(u)\, du = (-2i)^{-k}\{D_y^k e^{-iy^2}\Phi(\tau + y)\}_{y=0},$$

a homogeneous linear function of the derivatives $\Phi^{(l)}(\tau)(l = 0, \cdots, k)$ with numerical coefficients; moreover $z - \tau = \epsilon u + O(1)$ and consequently $\Phi^{(n)}(\tau) = O(1)$ $(n = 0, 1, \cdots)$. On the other hand, by Lemma 3, the expression $A_k\tau^{-k}$ is a polynomial in τ^{-1} of degree $\leqq k$ which does not contain the powers $\tau^{-\nu}$ for $0 \leqq \nu < \frac{k}{3}$. Therefore $\sum_{k=0}^{3n-1} A_kB_k\tau^{-k} = \sum_{k=0}^{3n-1} \Gamma_k^*\tau^{-k}$, where Γ_k^* is a homogeneous linear function of $\Phi^{(l)}(\tau)$ $(l = 0, \cdots, 3k)$ with constant coefficients and independent of n for $k < n$. Writing $\Gamma_k^* = \Gamma_k = \sum_{l=0}^{3k} \gamma_{kl}\Phi^{(l)}(\tau)$ $(k = 0, \cdots, n-1)$ and applying Lemma 4 with $3n$ instead of n, we obtain $F(s)/g(\tau) = \sum_{k=0}^{3n-1} A_kB_k\tau^{-k} + O(\tau^{-n}) = \sum_{k=0}^{n-1} \Gamma_k\tau^{-k} + O(\tau^{-n})$ $(n = 0, 1, \cdots)$, where $\Gamma_0 = \Phi(\tau)$; q.e.d.

2. Properties of the coefficients

In order to get simple recursion formulae for the coefficients γ_{kl}, we introduce the formal power series

$$d_l = d_l(T) = \sum_{k \geqq \frac{l}{3}} \gamma_{kl}\, T^{-k} \qquad (l = 0, 1, \cdots), \tag{20}$$

T being an indeterminate, and define $d_{-1} = d_{-2} = 0$.

LEMMA 5:

$$\frac{n(n+1)}{2}\, d_{n+1} + nTd_n + \tfrac{1}{8}d_{n-3} = 0 \qquad (n = 1, 2, \cdots).$$

PROOF: By (19),

$$\sum_{n=0}^{\infty} d_n\Phi^{(n)}(\tau) = \sum_{k=0}^{\infty} \Gamma_k\, T^{-k} = \sum_{k=0}^{\infty} A_k(T)B_k\, T^{-k}$$

$$= \sum_{k=0}^{\infty} A_k(T)(-2iT)^{-k}\{D_y^k e^{-iy^2}\Phi(\tau + y)\}_{y=0},$$

whence

$$d_n = \sum_{k=n}^{\infty} A_k(T)(-2iT)^{-k}\binom{k}{n}\{D^{k-n}e^{-iy^2}\}_{y=0}. \tag{21}$$

We write $A_k(T)(-2iT)^{-k}k! = G_k$, $\sum_{k=0}^{\infty} G_k y^{-k} = G$, with another indeterminate y; then the coefficients h_n of the formal Laurent series $L = e^{-iy^2}G = \sum_{n=-\infty}^{\infty} h_n y^{-n}$ are formal power series in T^{-1}, and by (21),

$$d_n = h_n/n! \qquad (n = 0, 1, \cdots). \tag{22}$$

Using Lemma 3, we obtain the recursion formula $TG_{n+1} + \frac{i}{2}\left(n + \frac{1}{2}\right)G_n +$

$\frac{n(n-1)}{8} G_{n-2} = 0$ $(n = 0, 1, \cdots)$; hence G satisfies formally the differential equation

$$T(G-1) - \frac{i}{2}\left(G' - \frac{1}{2y}G\right) + \frac{1}{8}(y^{-1}G)'' = 0.$$

Setting $G = e^{iy^2}L$, we find $(T + \frac{1}{2}y)L + \frac{1}{8}(y^{-1}L)'' = Te^{iy^2}$,

$$Th_n + \tfrac{1}{2}h_{n+1} + \frac{(n-1)(n-2)}{8} h_{n-3} = 0 \qquad (n = 1, 2, \cdots). \tag{23}$$

The assertion follows from (22), (23).

Let

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}}\left(\frac{\tau^2}{2}\right)^{\frac{1}{4}} e^{\frac{\pi}{4}\tau^2} \sqrt{\Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{1-s}{2}\right)} = \alpha = e^{\omega}, \\ &e^{-\frac{\pi i}{8}}\left(\frac{\tau^2}{2e}\right)^{\frac{i}{2}\tau^2} \sqrt{\Gamma\left(\frac{1-s}{2}\right)\Big/\Gamma\left(\frac{s}{2}\right)} = \beta = e^{-i\theta}. \end{aligned} \tag{24}$$

LEMMA 6:

$$\alpha\beta \approx d_0(\tau), \qquad \alpha^{-2} \approx 1 + \frac{d_1(\tau)}{d_0(\tau)}\tau^{-1}.$$

PROOF: In the special case $f(u) = 1$, we get

$$F(s) = F_0(s) = \int_{-\infty}^{+\infty} g(x)\,du = \pi 2^{\frac{1-s}{2}} e^{\frac{\pi i}{4}s} \Big/ \Gamma\left(\frac{s+1}{2}\right),$$

$$\Phi(y) = \Phi_0(y) = \int_{-\infty}^{+\infty} e^{i(x-y)^2}\,du = \sqrt{\pi};$$

consequently, by Theorem 1 and (20),

$$F_0(s)/g(\tau) \approx d_0(\tau)\Phi_0(\tau) = \sqrt{\pi}\, d_0(\tau).$$

For $f(u) = \epsilon u$, we obtain in the same manner

$$F(s) = F_1(s) = \int_{-\infty}^{+\infty} g(z)(z-p)\,du = F_0(s-1) - pF_0(s),$$

$$\Phi(y) = \Phi_1(y) = \int_{-\infty}^{+\infty} e^{i(x-y)^2}(z-p)\,du = (y-p)\int_{-\infty}^{+\infty} e^{i(x-y)^2}\,du = \sqrt{\pi}\,(y-p);$$

$$F_1(s)/g(\tau) \approx d_0(\tau)\Phi_1(\tau) + d_1(\tau)\Phi_1'(\tau) = \sqrt{\pi}\,\{(\tau-p)\,d_0(\tau) + d_1(\tau)\}.$$

Therefore

$$\frac{F_0(s-1)}{F_0(s)} - p \approx \tau - p + \frac{d_1(\tau)}{d_0(\tau)}, \qquad \tau^{-1}\frac{F_0(s-1)}{F_0(s)} \approx 1 + \frac{d_1(\tau)}{d_0(\tau)}\tau^{-1};$$

on the other hand

$$F_0(s-1)/F_0(s) = 2^{\frac{1}{2}}e^{-\frac{\pi i}{4}}\Gamma\left(\frac{s+1}{2}\right)\Big/\Gamma\left(\frac{s}{2}\right)$$
$$= \pi 2^{\frac{3}{2}}e^{\frac{\pi i}{2}\left(s-\frac{1}{2}\right)}\Big/(1+e^{\pi i s})\Gamma\left(\frac{s}{2}\right)\Gamma\left(\frac{1-s}{2}\right)$$
$$= \tau\alpha^{-2}/(1+e^{\pi i s}) = \tau\alpha^{-2}(1+O(e^{-\pi t})).$$

This proves $\alpha^{-2} \approx 1 + \tau^{-1}d_1(\tau)/d_0(\tau)$. Moreover

$$\pi^{-\frac{1}{2}}F_0(s)/g(\tau) = \pi^{\frac{1}{2}}2^{\frac{1-s}{2}}e^{\frac{\pi i}{4}s}\tau^s e^{-\frac{i}{2}\tau^2}\Big/\Gamma\left(\frac{s+1}{2}\right) = \alpha\beta(1+e^{\pi i s}),$$

whence $\alpha\beta \approx d_0(\tau)$; q.e.d.

The coefficients a_n, b_n of the power series

$$\frac{1}{\cos x} = \sum_{n=0}^{\infty}\frac{a_n}{(2n)!}x^{2n} \quad \left(|x| < \frac{\pi}{2}\right), \qquad \frac{x}{\sin x} = \sum_{n=0}^{\infty}\frac{b_n}{(2n!)}x^{2n} \quad (|x| < \pi)$$

may be calculated from the recursion formulae

$$\sum_{k=0}^{n}(-1)^k\binom{2n}{2k}a_k = 0, \qquad \sum_{k=0}^{n}(-1)^k\binom{2n+1}{2k}b_k = 0 \qquad (n = 1, 2, \cdots)$$

with $a_0 = b_0 = 1$.

LEMMA 7:

$$\omega \approx \frac{1}{8}\sum_{n=1}^{\infty}\frac{a_n}{n}(2\tau^2)^{-2n}, \qquad \Theta \approx \frac{1}{8}\sum_{n=1}^{\infty}\frac{b_n}{n(2n-1)}(2\tau^2)^{1-2n}.$$

PROOF: Let $\Re\{\xi\} \geqq \frac{1}{4}$, $\xi \neq \frac{1}{4}$; then

$$\frac{d\log\Gamma(\xi)}{d\xi} = \int_0^\infty\left(\frac{e^{-u}}{u} - \frac{e^{-\xi u}}{1-e^{-u}}\right)du = \int_0^\infty(e^{-u} - e^{(\frac{1}{4}-\xi)u})\frac{du}{u}$$
$$+ \int_0^\infty e^{(\frac{1}{4}-\xi)u}\left(u^{-1} - \frac{e^{-\frac{1}{4}u}}{1-e^{-u}}\right)du$$
$$= \log(\xi - \tfrac{1}{4}) - (4\xi-1)^{-1} + \int_0^\infty e^{(1-4\xi)u}\left(1 + u^{-1} - \frac{4e^{3u}}{e^{4u}-1}\right)du$$

$$\log\Gamma(\xi) = (\xi - \tfrac{1}{2})\log(\xi - \tfrac{1}{4}) - (\xi - \tfrac{1}{4})$$
$$+ \tfrac{1}{4}\int_0^\infty e^{(1-4\xi)u}\left(\frac{1}{\mathrm{ch}u} + \frac{1}{\mathrm{sh}u} - u^{-1} - 1\right)\frac{du}{u} + c$$

with constant c. On account of the formula $\Gamma(\xi)\Gamma(\xi+\frac{1}{2}) = \sqrt{\pi}\,2^{1-2\xi}\Gamma(2\xi)$, the passage to the limit $\xi \to \infty$ gives the value $c = \frac{1}{2}\log 2\pi$. Applying Cauchy's theorem, we obtain

$$\log\Gamma\left(\frac{s}{2}\right) = \frac{s-1}{2}\log\left(\frac{s}{2} - \frac{1}{4}\right) - \left(\frac{s}{2} - \frac{1}{4}\right) + \frac{1}{2}\log 2\pi$$

$$+\frac{1}{4}\int_0^\infty e^{i(2s-1)x}\left\{\frac{1}{\cos x}-1+i\left(\frac{1}{\sin x}-x^{-1}\right)\right\}\frac{dx}{x},$$

where the poles $x=\frac{k\pi}{2}$ $(k=1, 2, \cdots)$ of the integrand on the positive real axis are avoided by small half-circles to the left; hence

$$\log\Gamma\left(\frac{s}{2}\right)-\frac{s-1}{2}\log\left(\frac{s}{2}-\frac{1}{4}\right)$$

$$+\left(\frac{s}{2}-\frac{1}{4}\right)-\frac{1}{2}\log 2\pi=\frac{1}{8}\sum_{n=1}^{\infty}\left\{\frac{a_n}{n}(2\tau^2)^{-2n}+i\frac{b_n}{n(2n-1)}(2\tau^2)^{1-2n}\right\},$$

and the assertion follows from the definition of ω, Θ in (24).

THEOREM 2:

$$e^{i\Theta}F(s)/g(\tau)\approx\sum_{k=0}^{\infty}G_k\tau^{-k},\qquad G_k=\sum_{l=0}^{3k}g_{kl}\Phi^{(l)}(\tau), \tag{25}$$

where the coefficients g_{kl} are rational numbers computed from the recursion formulae

$$\delta_{-2}=\delta_{-1}=0,\qquad \delta_0=e^{\omega(T)},\qquad \delta_1=(\delta_0^{-1}-\delta_0)T, \tag{26}$$

$$\delta_n=-\frac{2T}{n}\delta_{n-1}-\frac{1}{4n(n-1)}\delta_{n-4}\qquad (n=2, 3, \cdots), \tag{27}$$

$$\omega(T)=\tfrac{1}{8}\sum_{n=1}^{\infty}\frac{a_n}{n}(2T^2)^{-2n},\qquad a_0=1,$$

$$\sum_{k=0}^{n}(-1)^k\binom{2n}{2k}a_k=0\qquad (n=1, 2, \cdots), \tag{28}$$

$$\delta_n=\sum_{k\geqq\frac{n}{3}}g_{kn}T^{-k}\qquad (n=0, 1, \cdots). \tag{29}$$

PROOF: Defining $\Theta(T)=\frac{1}{8}\sum_{n=1}^{\infty}\frac{b_n}{n(2n-1)}(2T^2)^{1-2n}$, $\delta_n=e^{i\Theta(T)}d_n (n=0, 1, \cdots)$, we infer from Lemmata 5, 6, 7 that

$$\delta_0=e^{\omega(T)},\qquad \delta_0^{-2}=1+\frac{\delta_1}{\delta_0}T^{-1},$$

$$\frac{n(n+1)}{2}\delta_{n+1}+nT\delta_n+\tfrac{1}{8}\delta_{n-3}=0\qquad (n=1, 2, \cdots),$$

i.e. (26), (27), (28). Formulae (25), (29) follow from Theorem 1, (20), Lemma 7. By (26), (27), (28), (29), all coefficients g_{kl} are rational numbers.

We find in particular

$$G_0=\Phi(\tau),\qquad G_1=-\frac{1}{2^3.3}\Phi^{(3)}(\tau),\qquad G_2=\frac{1}{2^4}\Phi^{(2)}(\tau)+\frac{1}{2^7.3^2}\Phi^{(6)}(\tau),$$

$$G_3=-\frac{1}{2^4}\Phi^{(1)}(\tau)-\frac{1}{2^4.3.5}\Phi^{(5)}(\tau)-\frac{1}{2^{10}.3^4}\Phi^{(9)}(\tau),$$

$$G_4 = \frac{1}{2^5}\Phi(\tau) + \frac{19}{2^9.3}\Phi^{(4)}(\tau) + \frac{11}{2^{11}.3^2.5}\Phi^{(8)}(\tau) + \frac{1}{2^{15}.3^5}\Phi^{(12)}(\tau).$$

We shall use the symbol $\Omega(\Phi)$ to designate the asymptotic series $\sum_{k=0}^{\infty} G_k \tau^{-k}$.

3. Calculation of some integrals

By the sign $l \swarrow l+1$ we mean that the integration extends over a line $x = \epsilon u + a$, where a is a given real number in the interval $l < a < l+1$, $\epsilon = e^{\frac{\pi i}{4}}$ and u runs through real values from $+\infty$ to $-\infty$. Let q be a positive integer and ξ a complex parameter.

LEMMA 8:

$$\frac{1}{2i}\int_{-1\swarrow 0} e^{q\pi i(x-\xi)^2} \operatorname{ctg} \pi x \, dx = \frac{e^{q\pi i\xi^2}}{1 - e^{q\pi i(1-2\xi)}} + \sum_{k=1}^{q} \alpha_k \operatorname{tg} \pi\left(\xi + \frac{k}{q}\right)$$

with constant $\alpha_1, \cdots, \alpha_q$.

PROOF: We denote the integral by J and apply Cauchy's theorem; then

$$\begin{aligned} J - e^{q\pi i\xi^2} &= \frac{1}{2i}\int_{0\swarrow 1} e^{q\pi i(x-\xi)^2} \operatorname{ctg} \pi x \, dx \\ &= \frac{1}{2i} e^{q\pi i(1-2\xi)} \int_{1\swarrow 0} e^{q\pi i(x-\xi)^2 + 2q\pi i x} \operatorname{ctg} \pi x \, dx \\ &= e^{q\pi i(1-2\xi)} \left\{ J + \frac{1}{2i}\int_{1\swarrow 0} e^{q\pi i(x-\xi)^2}(e^{2q\pi i x} - 1) \operatorname{ctg} \pi x \, dx \right\}. \end{aligned}$$

Since

$$\frac{1}{2i}(e^{2q\pi i x} - 1) \operatorname{ctg} \pi x = \sum_{k=0}^{q} \beta_k e^{2k\pi i x}, \quad \beta_0 = \beta_q = \tfrac{1}{2}, \quad \beta_k = 1 \qquad (0 < k < q)$$

and

$$\frac{1}{2i}\int_{1\swarrow 0} e^{q\pi i(x-\xi)^2 + 2k\pi i x} dx = e^{-\pi i\frac{k^2}{q} + 2k\pi i\xi} b,$$

where

$$b = \int_{-1\swarrow 0} e^{q\pi i\left(x-\xi+\frac{k}{q}\right)^2} dx = \int_{-1\swarrow 0} e^{q\pi i x^2} dx = -\epsilon q^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-\pi x^2} dx \tag{30}$$

is independent of ξ, we get

$$(1 - \eta^q)J = e^{q\pi i\xi^2} + \sum_{k=0}^{q} \beta_k^* \eta^{q-k} \tag{31}$$

with

$$\eta = e^{\pi i(1-2\xi)}, \qquad \beta_k^* = (-1)^k e^{-\pi i\frac{k^2}{q}} \beta_k b \qquad (k = 0, \cdots, q). \tag{32}$$

On the other hand,

$$(33)\quad \sum_{k=0}^{q} \beta_k^* \eta^{q-k}/(1-\eta^q) = \alpha_0 + i\sum_{k=1}^{q} \alpha_k \frac{\eta+\epsilon_k}{\eta-\epsilon_k} = \alpha_0 + \sum_{k=1}^{q} \alpha_k \operatorname{tg} \pi\left(\xi + \frac{k}{q}\right),$$

where $\alpha_0, \alpha_1, \cdots, \alpha_q$ are certain constants and $\epsilon_k = e^{2\pi i \frac{k}{q}}$. Choosing $\eta = 0$ and $\eta = \infty$, we find $2\alpha_0 = \beta_q^* - \beta_0^* = (\beta_q - \beta_0)\, b = 0$, and the assertion follows from (31), (33).

LEMMA 9:

$$\sum_{k=1}^{q} e^{-\pi i \frac{k^2}{q} + \pi i \frac{kn}{q}} = e^{\frac{\pi i}{4}\left(\frac{n^2}{q}-1\right)} q^{\frac{1}{2}} \qquad (n-q \text{ even}).$$

PROOF: Take $\xi = \dfrac{n}{2q}$ and apply (31), (32); then

$$\eta = e^{\pi i\left(1-\frac{n}{q}\right)}, \qquad \eta^q = 1, \qquad b\sum_{k=1}^{q} e^{-\pi i \frac{k^2}{q} + \pi i \frac{kn}{q}} = -e^{\pi i \frac{n^2}{4q}},$$

and the assertion follows from (30).

We introduce a proper character $\chi(n)$ modulo $m \geqq 1$ and define

$$W(x) = \frac{\pi}{2m}\sum_{n=1}^{2m} \chi(n) e^{-\pi i \frac{n^2}{m}} \operatorname{ctg} \pi \frac{x-n}{2m}, \qquad C = C(\chi) = \sum_{n=1}^{m} \chi(n) e^{-2\pi i \frac{n}{m}}.$$

THEOREM 3:

$$\int_{0\swarrow 1} e^{\frac{\pi i}{m}x^2 - \frac{2\pi i}{m}\xi x} W(x)\, dx = \frac{2\pi i}{1-e^{-2\pi i\xi}} \sum_{n=1}^{m} \chi(n) e^{-\frac{2\pi i n}{m}\xi} - Ce^{-\frac{\pi i}{m}\xi^2}\overline{W(\bar{\xi})}.$$

PROOF: Apply Lemma 8 with $q = 4m$, replace x, ξ by $\dfrac{x-n}{2m}, \dfrac{\xi-n}{2m}$, multiply by $2\pi i \chi(n) e^{-\pi i \frac{n^2}{m}}$ and sum over n from 1 to $2m$; then

$$\int_{0\swarrow 1} e^{\frac{\pi i}{m}(x-\xi)^2} W(x)\, dx = \frac{2\pi i}{1-e^{-4\pi i\xi}} \sum_{n=1}^{2m} \chi(n) e^{-\pi i \frac{n^2}{m} + \frac{\pi i}{m}(\xi-n)^2}$$

$$(34)\qquad + 2\pi i \sum_{n=1}^{2m} \chi(n) e^{-\pi i \frac{n^2}{m}} \sum_{k=1}^{4m} \alpha_k \operatorname{tg} \frac{\pi}{2m}\left(\xi - n + \frac{k}{2}\right)$$

$$= \frac{2\pi i e^{\frac{\pi i}{m}\xi^2}}{1-e^{-2\pi i\xi}} \sum_{n=1}^{m} \chi(n) e^{-\frac{2\pi i n}{m}\xi} + \sum_{k=1}^{4m} \lambda_k \operatorname{ctg} \frac{\pi}{2m}\left(\xi - \frac{k}{2}\right)$$

with constant $\lambda_1, \cdots, \lambda_{4m}$. Performing the passage to the limit $\xi \to \dfrac{k}{2}$, we find $\lambda_k = 0$ (k odd) and

$$(35)\qquad \frac{2m}{\pi}\lambda_{2l} = -e^{\pi i \frac{l^2}{m}} \sum_{n=1}^{m} \chi(n) e^{-2\pi i \frac{ln}{m}} \qquad (l = 1, \cdots, 2m).$$

Since

$$\sum_{n=1}^{m} \chi(n) e^{-2\pi i \frac{ln}{m}} = C\bar{\chi}(l), \tag{36}$$

the assertion follows from (34), (35).

We set

$$\gamma_k = 0 \quad (k - m \text{ odd}), \qquad \gamma_k = \frac{i}{8m} e^{\pi i \frac{k^2}{8m}} \sum_{l=1}^{m} \chi(l) e^{\pi i \frac{l^2}{m} - \pi i \frac{kl}{m}} \quad (k - m \text{ even}).$$

LEMMA 10:

$$C\bar{\gamma}_k = -e^{-\frac{\pi i}{4}} m^{\frac{1}{2}} \gamma_k \qquad (k = 1, \cdots, 8m).$$

PROOF: The assertion is trivial if $k - m$ is odd. In the other case, by Lemma 9 and (36),

$$\begin{aligned} 8miCe^{\pi i \frac{k^2}{8m}} \bar{\gamma}_k &= \sum_{l,n=1}^{m} \chi(n) e^{-2\pi i \frac{ln}{m} - \pi i \frac{l^2}{m} + \pi i \frac{kl}{m}} \\ &= \sum_{n=1}^{m} \chi(n) e^{\pi i \frac{n^2}{m} - \pi i \frac{kn}{m}} \sum_{l=1}^{m} e^{-\pi i \frac{l^2}{m} + \pi i \frac{kl}{m}} \\ &= -8mie^{-\pi i \frac{k^2}{8m}} \gamma_k e^{\frac{\pi i}{4}\left(\frac{k^2}{m} - 1\right)} m^{\frac{1}{2}}; \end{aligned}$$

q.e.d.

LEMMA 11: If l *is divisible by* m, *then*

$$\begin{aligned} &\frac{1}{2\pi i} \int_{l \swarrow l+1} e^{\frac{2\pi i}{m}(x-\xi)^2} W(x)\, dx \\ &\quad = \frac{(-1)^l}{1 - (-1)^m e^{-4\pi i \xi}} \sum_{n=1}^{m} \chi(n) e^{-\pi i \frac{n^2}{m} + \frac{2\pi i}{m}(\xi - l - n)^2} + \sum_{k=1}^{8m} \gamma_k \operatorname{ctg} \frac{\pi}{2m}\left(\xi - \frac{k}{4}\right). \end{aligned}$$

PROOF: Apply Lemma 8 with $q = 8m$, replace x, ξ by $\frac{x - n}{2m}$, $\frac{\xi - n - l}{2m}$, multiply by $\chi(n) e^{-\pi i \frac{n^2}{m}}$ and sum over n from 1 to $2m$; then

$$\begin{aligned} \frac{1}{2\pi i} \int_{0 \swarrow 1} e^{\frac{2\pi i}{m}(x+l-\xi)^2} W(x)\, dx &= \frac{1}{1 - e^{-8\pi i \xi}} \sum_{n=1}^{2m} \chi(n) e^{-\pi i \frac{n^2}{m} + \frac{2\pi i}{m}(\xi - n - l)^2} \\ &+ \sum_{n=1}^{2m} \chi(n) e^{-\pi i \frac{n^2}{m}} \sum_{k=1}^{8m} \alpha_k \operatorname{tg} \frac{\pi}{2m}\left(\xi - n - l + \frac{k}{4}\right) \qquad (37) \\ &= \frac{e^{\frac{2\pi i}{m}(\xi - l)^2}}{1 - (-1)^m e^{-4\pi i \xi}} \sum_{n=1}^{m} \chi(n) e^{\pi i \frac{n^2}{m} - \frac{4\pi i n}{m} \xi} + \sum_{k=1}^{8m} \gamma_k^* \operatorname{ctg} \frac{\pi}{2m}\left(\xi - \frac{k}{4}\right) \end{aligned}$$

with constant $\gamma_1^*, \cdots, \gamma_{8m}^*$. Performing the passage to the limit $\xi \to \frac{k}{4}$, we

find $\gamma_k^* = 0$ ($k - m$ odd) and

$$(38) \qquad \gamma_k^* = \frac{i}{8m} e^{\frac{2\pi i}{m}\left(\frac{k}{4}-l\right)^2} \sum_{n=1}^{m} \chi(n) e^{\pi i \frac{n^2}{m} - \pi i \frac{nk}{m}} = (-1)^l \gamma_k \qquad (k - m \text{ even}).$$

Since $W(x + l) = (-1)^l W(x)$, the assertion follows from (37), (38).

4. Generalization of Riemann's formulae

It is well-known that C has the absolute value $m^{\frac{1}{2}}$. We define

$$a = \frac{1 - \chi(-1)}{2}, \qquad \rho = i^{-\frac{a}{2}} C^{-\frac{1}{2}} m^{\frac{1}{4}}, \qquad \lambda(s) = \frac{1}{2\pi i} \int_{0\swarrow 1} e^{\frac{\pi i}{m} x^2} x^{-s} W(x)\, dx,$$

$$\mu(s) = \rho \left(\frac{m}{\pi}\right)^{\frac{s}{2}} \Gamma\left(\frac{s+a}{2}\right) \lambda(s), \qquad L(s) = \sum_{n=1}^{\infty} \chi(n) n^{-s} \qquad (\sigma > 1).$$

THEOREM 4:

$$\rho \left(\frac{m}{\pi}\right)^{\frac{s}{2}} \Gamma\left(\frac{s+a}{2}\right) L(s) = \mu(s) + \overline{\mu(1 - \bar{s})}.$$

PROOF:

$$\int_0^{\epsilon^{-1}\infty} \frac{\xi^{s-1}}{1 - e^{-2\pi i \xi}} \sum_{n=1}^{m} \chi(n) e^{-\frac{2\pi i n}{m}\xi}\, d\xi = \int_0^{\epsilon^{-1}\infty} \xi^{s-1} \sum_{n=1}^{\infty} \chi(n) e^{-\frac{2\pi i n}{m}\xi}\, d\xi$$

$$= \Gamma(s) \sum_{n=1}^{\infty} \chi(n) \left(\frac{2\pi i n}{m}\right)^{-s} = \left(\frac{m}{2\pi}\right)^{s} e^{-\frac{\pi i}{2}s} \Gamma(s) L(s),$$

$$\int_0^{\epsilon^{-1}\infty} \xi^{s-1} \left(\int_{0\swarrow 1} e^{\frac{\pi i}{m}x^2 - \frac{2\pi i}{m}\xi x} W(x)\, dx \right) d\xi$$

$$= \left(\frac{m}{2\pi}\right)^{s} e^{-\frac{\pi i}{2}s} \Gamma(s) \int_{0\swarrow 1} e^{\frac{\pi i}{m}x^2} x^{-s} W(x)\, dx;$$

consequently, by Theorem 3 and the formula $W(-x) = -\chi(-1)W(x)$,

$$\left(\frac{m}{2\pi}\right)^{s} e^{-\frac{\pi i}{2}s} \Gamma(s) \{L(s) - \lambda(s)\} = \frac{C}{2\pi i} \int_0^{\epsilon^{-1}\infty} e^{-\frac{\pi i}{m}\xi^2} \xi^{s-1} \overline{W(\bar{\xi})}\, d\xi$$

$$= -\frac{C}{1 + \chi(-1)e^{\pi i s}} \frac{1}{2\pi i} \int_{0\nwarrow 1} e^{-\frac{\pi i}{m}\xi^2} \xi^{s-1} \overline{W(\bar{\xi})}\, d\xi = \frac{C\overline{\lambda(1 - \bar{s})}}{1 + \chi(-1)e^{\pi i s}};$$

moreover

$$2^s e^{\frac{\pi i}{2}s} \Gamma\left(\frac{s+a}{2}\right) \Big/ \Gamma(s) = \pi^{-\frac{1}{2}} i^a (1 + \chi(-1)e^{\pi i s}) \Gamma\left(\frac{1 - s + a}{2}\right),$$

and the assertion follows.

Set

$$l = m\left[\sqrt{\frac{t}{2\pi m}}\right], \qquad \eta = y\sqrt{\frac{m}{2\pi}} - l - \frac{1}{2},$$

$$\Phi(y) = \frac{1}{1-(-1)^m e^{-4\pi i\eta}} \sum_{n=1}^{m} \chi(n) e^{-\pi i \frac{n^2}{m} + \frac{2\pi i}{m}\left(\eta+\frac{1}{2}-n\right)^2} + (-1)^l \sum_{k=1}^{8m} \gamma_k \operatorname{ctg} \frac{\pi}{2m}\left(\eta + l + \frac{1}{2} - \frac{k}{4}\right).$$

THEOREM 5:

$$\lambda(s) = \sum_{n=1}^{l} \chi(n) n^{-s} + (-1)^l e^{\frac{i}{2}\tau^2 - i\Theta} \left(\frac{2\pi}{m}\right)^{\frac{s}{2}} \tau^{-s} Q, \qquad Q \approx \Omega(\Phi).$$

PROOF: Putting $\sqrt{\frac{m}{2\pi}} = b$, $p = (l+\frac{1}{2})b^{-1}$, $z = \epsilon u + p$, $x = bz$ and $f(u) = -\frac{\epsilon b}{2\pi i} W(\frac{1}{2} + \epsilon b u)$, we infer from Cauchy's theorem that

$$\lambda(s) - \sum_{n=1}^{l} \chi(n) n^{-s} = \frac{1}{2\pi i} \int_{l \swarrow l+1} e^{\frac{\pi i}{m} x^2} x^{-s} W(x)\, dx = (-1)^l b^{-s} \int_{-\infty}^{+\infty} g(z) f(u)\, du.$$

By Lemma 11

$$\int_{-\infty}^{+\infty} e^{i(z-y)^2} f(u)\, du = \frac{(-1)^l}{2\pi i} \int_{l \swarrow l+1} e^{\frac{2\pi i}{m}(x-by)^2} W(x)\, dx$$

$$= \frac{1}{1-(-1)^m e^{-4\pi i b y}} \sum_{n=1}^{m} \chi(n) e^{-\pi i \frac{n^2}{m} + \frac{2\pi i}{m}(by-l-n)^2} + (-1)^l \sum_{k=1}^{8m} \gamma_k \operatorname{ctg} \frac{\pi}{2m}\left(by - \frac{k}{4}\right) = \Phi(y),$$

and the assertion follows from Theorem 2.

We define $\chi(n) = e^{i\alpha_n}$, if n and m are coprime, $\rho e^{\frac{\pi i}{8}(2a-1)} = e^{i\delta}$,

$$\Psi(y) = \frac{(-1)^{l-1}}{\sin 2\pi\left(\eta + \frac{m}{4}\right)} \sum_{\substack{n=1 \\ (n,m)=1}}^{m} \sin\left\{\frac{2\pi}{m}\left(\eta + \frac{m+1}{2} - n\right)^2 - \pi\frac{n^2}{m} + \alpha_n + \delta\right\},$$

$$e^{i\vartheta} = \rho\left(\frac{m}{\pi}\right)^{\frac{s}{2}-\frac{1}{4}} \sqrt{\Gamma\left(\frac{s+a}{2}\right) \Big/ \Gamma\left(\frac{1-s+a}{2}\right)}.$$

THEOREM 6:

$$e^{i\vartheta} L(s) = 2 \sum_{\substack{n=1 \\ (n,m)=1}}^{l} n^{-\frac{1}{2}} \cos(\vartheta + \alpha_n - \tau^2 \log n) + \left(\frac{2\pi}{m}\right)^{\frac{1}{2}} \tau^{-\frac{1}{2}} R, \qquad R \approx \Omega(\Psi).$$

PROOF: By (24),

$$(39) \quad e^{\frac{i}{2}\tau^2 - i\Theta} \left(\frac{2\pi}{m}\right)^{\frac{s}{2}-\frac{1}{4}} \tau^{\frac{1}{2}-s} = e^{-\frac{\pi i}{8}} \left(\frac{\pi}{m}\right)^{\frac{s}{2}-\frac{1}{4}} \sqrt{\Gamma\left(\frac{1-s}{2}\right) \Big/ \Gamma\left(\frac{s}{2}\right)} = \rho e^{-\frac{\pi i}{8} - i\vartheta} c,$$

where $c = 1$ for $a = 0$ and $c = \operatorname{tg}^{\frac{1}{2}} \frac{\pi s}{2}$ for $a = 1$ and consequently $c = i^{\frac{a}{2}} +$

$O(e^{-\pi t})$. The functions

$$\vartheta = \vartheta(s) \text{ and } \nu(s) = \rho\left(\frac{m}{\pi}\right)^{\frac{s}{2}}\Gamma\left(\frac{s+a}{2}\right)e^{-i\vartheta} = \left(\frac{\pi}{m}\right)^{\frac{1}{4}}\sqrt{\Gamma\left(\frac{s+a}{2}\right)\Gamma\left(\frac{1-s+a}{2}\right)}$$

satisfy $\vartheta(s) = -\overline{\vartheta(1-\bar{s})}$ and $\nu(s) = \overline{\nu(1-\bar{s})}$. It follows from Theorems 4, 5 and (39) that

$$e^{i\vartheta}L(s) = e^{i\vartheta}\lambda(s) + e^{-i\vartheta}\overline{\lambda(1-\bar{s})}$$

$$= 2\sum_{\substack{n=1\\(n,m)=1}}^{l} n^{-\frac{1}{2}}\cos(\vartheta + \alpha_n - \tau^2\log n) + \left(\frac{2\pi}{m}\right)^{\frac{1}{2}}\tau^{-\frac{1}{2}}R,$$

$$R \ll \Omega(\psi), \qquad (-1)^l\psi(y) = \rho e^{\frac{\pi i}{8}(2a-1)}\Phi(y) + \rho^{-1}e^{-\frac{\pi i}{8}(2a-1)}\overline{\Phi(\bar{y})}.$$

By Lemma 10,

$$\rho e^{\frac{\pi i}{8}(2a-1)}\gamma_k + \rho^{-1}e^{-\frac{\pi i}{8}(2a-1)}\overline{\gamma_k} = \rho e^{\frac{\pi i}{8}(2a-1)}\left(\gamma_k + e^{\frac{\pi i}{4}}Cm^{-\frac{1}{2}}\overline{\gamma_k}\right) = 0;$$

hence

$$\psi(y) = \frac{(-1)^l}{\sin 2\pi\left(\eta + \frac{m}{4}\right)}\sum_{\substack{n=1\\(n,m)=1}}^{m}$$

$$\sin\left\{\frac{2\pi}{m}(\eta + \tfrac{1}{2} - n)^2 - \pi\frac{n^2}{m} + 2\pi\eta + \pi\frac{m}{2} + \alpha_n + \delta\right\} = \Psi(y);$$

q.e.d.

For our further purposes we need Theorem 4 and the main term in Theorem 5, namely the formula

$$\lambda(s) = \sum_{n=1}^{r}\chi(n)n^{-s} + O\left(t^{-\frac{\sigma}{2}}\right), \qquad r = \left[\sqrt{\frac{mt}{2\pi}}\right]; \tag{40}$$

obviously

$$\log\lambda(s) = \log\left(L(s) + O(t^{-1})\right) = \sum_{k,p}\frac{1}{k}\chi(p^k)p^{-ks} + O(t^{-1}) \qquad (\sigma = 3), \tag{41}$$

p running over all prime numbers and k over all positive integers. Moreover, by Lemma 7 and (39),

$$\Re\{\vartheta\} = \Re\left\{\frac{\tau^2}{2}\log\frac{m\tau^2}{2\pi e}\right\} + O(1) = \frac{t}{2}\log\frac{mt}{2\pi e} + O(1), \tag{42}$$

$$|e^{i\vartheta}| = \left(\frac{mt}{2\pi}\right)^{\frac{\sigma}{2}-\frac{1}{4}}(1 + O(t^{-1})) = O\left(t^{\frac{\sigma}{2}-\frac{1}{4}}\right). \tag{43}$$

5. The zeros of $L(s)$ on and near the critical line

LEMMA 12: *Let $h(s)$ be a regular analytic function on a segment C and δ the variation of* arg $h(s)$ *on C, where the zeros of $h(s)$ on C are avoided by small half-*

circles to the left; let N be the number of different points on C, where the real part of $h(s)$ changes its sign; then

$$N \geqq \frac{\delta}{\pi} - 1. \tag{44}$$

PROOF: Assume first that $h(s) \neq 0$ on C, then $\arg h(s)$ is continuous on C and runs over an interval of length $\geqq |\delta|$. The number of odd multiples of $\frac{\pi}{2}$ in the interior of this interval is $\geqq \frac{|\delta|}{\pi} - 1$; hence $N \geqq \frac{|\delta|}{\pi} - 1 \geqq \frac{\delta}{\pi} - 1$.

Let now $s = s_k (k = 1, \cdots, n)$ be all zeros of $h(s)$ on C, every zero written with its multiplicity; let α be the angle between C and the real axis, $a = e^{i\alpha}$; then $h_1(s) = h(s) \prod_{k=1}^{n} \left(\frac{a}{s - s_k}\right)$ is regular and $\neq 0$ on C, $\Re\{h(s)\} = \Re\{h_1(s)\} \prod_{k=1}^{n} \left(\frac{s - s_k}{a}\right)$. Denoting by δ_1 the variation of $\arg h_1(s)$ on C, we have $\delta_1 = \delta + \pi n$ and therefore, if the real part of $h_1(s)$ changes its sign exactly N_1 times, $N_1 \geqq \frac{\delta_1}{\pi} - 1 = \frac{\delta}{\pi} + n - 1$. On the other hand, $N \geqq N_1 - n$, and (44) follows.

We define $\beta(s) = e^{i\vartheta}\lambda(s) = \rho \left(\frac{m}{\pi}\right)^{\frac{s}{2} - \frac{1}{4}} \sqrt{\Gamma\left(\frac{s+a}{2}\right) \Big/ \Gamma\left(\frac{1-s+a}{2}\right)}\,\lambda(s)$; by Theorem 4,

$$\tfrac{1}{2} e^{i\vartheta} L(s) = \Re\{\beta(s)\} \qquad (\sigma = \tfrac{1}{2}). \tag{45}$$

LEMMA 13: *Let $\delta(t)$, $\delta_1(t)$ be the variations of* $\arg \beta(s)$, $\arg \lambda(s)$ *on the segment $s = \sigma + ti$, the real part σ running increasingly or decreasingly over a given interval $\sigma_3 \leqq \sigma \leqq \sigma_4$; then $\delta(t) = O(\log t)$, $\delta_1(t) = O(\log t)$.*

PROOF: Consider the function $\gamma(z) = \frac{1}{2}\lambda(z) + \frac{1}{2}\overline{\lambda(2ti + \bar{z})}$ in the circle $|z - z_0| \leqq 2r$, $z_0 = 2 + ti$, $r = 2 + |\sigma_3| + |\sigma_4|$. By (40),

$$|\gamma(z)| < \sum_{n < c_{16} t^{\frac{1}{2}}} n^{2r-2} + O(t^{r-1}) = O(t^{r-\frac{1}{2}}),$$

$$\gamma(z_0) = \Re\{\lambda(z_0)\} > 2 - \zeta(2) + O(t^{-\frac{1}{2}}) > c_{17} \qquad (t > c_{18}).$$

On account of Jensen's theorem, the number of zeros of $\gamma(z)$ in the circle $|z - z_0| \leqq r$ is therefore $O(\log t)$. This circle contains the segment $z = \sigma + ti$, $\sigma_3 \leqq \sigma \leqq \sigma_4$, and $\gamma(\sigma + ti) = \Re\{\lambda(\sigma + ti)\}$; consequently the number of zeros of $\Re\{\lambda(\sigma + ti)\}$ on the segment is $O(\log t)$. By Lemma 12, the variation $\delta_1(t)$ of $\arg \lambda(s)$ is $O(\log t)$. By (42), the variation $\delta_2(t)$ of $\arg e^{i\vartheta} = \Re\{\vartheta\}$ is $O(1)$. Since $\delta(t) = \delta_1(t) + \delta_2(t)$, the assertion follows.

We denote by $A(t_1, t_2)$ the number of zeros of odd order of the function $L(\frac{1}{2} + ui)$ in the interval $t_1 < u < t_2$, where $0 < t_1 < t_2$.

LEMMA 14: *Let σ_0 be any number in the interval $\sigma_1 \leqq \sigma < \frac{1}{2}$; then*

$$\pi(\tfrac{1}{2} - \sigma_0) A(t_1, t) > -\int_{t_1}^{t} \log |\beta(\sigma_0 + ui)|\, du + O(\log t).$$

PROOF: Assume that a function $h(s)$ is regular in the rectangle $\sigma_1 \leqq \sigma \leqq \sigma_2$, $t_1 \leqq t \leqq t_2$. We define $\arg h(s)$ on the contour by analytic continuation, beginning at the point $s_1 = \sigma_1 + t_1 i$ and running through the boundary C in positive direction, where zeros of $h(s)$ are avoided by small half-circles to the left. Then

$$\frac{1}{2\pi i}\int_C \log h(s)\, ds = \sum_\rho (s_1 - \rho),$$

where ρ runs over all zeros of $h(s)$ in the interior, whence

$$-\Im\left\{\int_C \log h(s)\, ds\right\} = 2\pi \sum_\rho (\Re\{\rho\} - \sigma_1) \geqq 0,$$

(46)
$$\int_{t_1}^{t_2} \log |\, h(\sigma_1 + ti)\,|\, dt + \int_{\sigma_1}^{\sigma_2} \arg h(\sigma + t_2 i)\, d\sigma \geqq \int_{t_1}^{t_2} \log |\, h(\sigma_2 + ti)\,|\, dt + \int_{\sigma_1}^{\sigma_2} \arg h(\sigma + t_1 i)\, d\sigma.$$

Applying this inequality with $h(s) = \beta(s)$, $\sigma_1 = \sigma_0$, $\sigma_2 = \frac{1}{2}$ and with $h(s) = \lambda(s)$, $\sigma_1 = \frac{1}{2}$, $\sigma_2 = 3$, we obtain, by (41) and Lemma 13,

(47)
$$\int_{t_1}^{t} \log |\, \beta(\sigma_0 + ui)\,|\, du + \int_{\sigma_0}^{\frac{1}{2}} \arg \beta(\sigma + ti)\, d\sigma > \int_{t_1}^{t} \log |\, \beta(\tfrac{1}{2} + ui)\,|\, du + O(\log t),$$
$$\int_{t_1}^{t} \log |\, \lambda(\tfrac{1}{2} + ui)\,|\, du > O(\log t).$$

Let δ be the variation of $\arg \beta(\frac{1}{2} + ui)$ for $t_1 \leqq u \leqq t$; then, by (45) and Lemmata (12), (13),

(48) $$\arg \beta(\sigma + ti) = \delta + O(\log t) < \pi A(t_1, t) + O(\log t) \quad (\sigma_0 \leqq \sigma \leqq \tfrac{1}{2});$$

moreover $|\, \beta(\frac{1}{2} + ui)| = |\, \lambda(\frac{1}{2} + ui)|$. This proves

$$\int_{t_1}^{t} \log |\, \beta(\sigma_0 + ui)\,|\, du + \pi(\tfrac{1}{2} - \sigma_0)\, A(t_1, t) > O(\log t);$$

q.e.d.

LEMMA 15: *If* $\sigma_1 \leqq \sigma < \frac{1}{2}$ *and* $t^{\frac{1}{2}} \log t = o(t - t_1)$, *then*

(39) $$\int_{t_1}^{t} \log |\, \beta(\sigma + ui)\,|\, du < \frac{t - t_1}{2}\left(\log \frac{\varphi(m)}{m(1 - 2\sigma)} + o(1)\right).$$

PROOF: Let $t - t_1 = \Delta$, $r = r(t) = \left[\sqrt{\frac{mt}{2\pi}}\right]$, $\sum_{n=1}^{r} \chi(n) n^{-s} = \lambda_0(s)$, then

(50) $$\int_{t_1}^{t} \log |\, \beta(\sigma + ui)\,|\, du \leqq \Delta \log\left(\Delta^{-1} \int_{t_1}^{t} |\, \beta(\sigma + ui)\,|\, du\right),$$

and by (40), (43),

$$(51)\qquad \int_{t_1}^{t} |\beta(\sigma+ui)|\,du < \int_{t_1}^{t} \left(\frac{mu}{2\pi}\right)^{\frac{\sigma}{2}-\frac{1}{4}} |\lambda_0(\sigma+ui)|\,du + \Delta O(t^{-\frac{1}{4}}).$$

Suppose first $t < 2t_1$ and set

$$\int_{t_1}^{t} \left(\frac{mu}{2\pi}\right)^{\frac{\sigma}{2}-\frac{1}{4}} |\lambda_0(\sigma+ui)|\,du = H.$$

Since

$$\sum_{n=1}^{r(u)} |\chi(n)|^2 n^{-2\sigma} < \frac{\varphi(m)}{m}\int_0^{r(u)} v^{-2\sigma}\,dv + O(1) = \frac{\varphi(m)}{m(1-2\sigma)}\left(\frac{mu}{2\pi}\right)^{\frac{1}{2}-\sigma}(1+o(1)),$$

we obtain

$$\Delta^{-2}H^2 \leqq \Delta^{-1}\int_{t_1}^{t}\left(\frac{mu}{2\pi}\right)^{\sigma-\frac{1}{2}} |\lambda_0(\sigma+ui)|^2\,du < \frac{\varphi(m)}{m(1-2\sigma)}(1+o(1))$$
$$+\Delta^{-1}O\left(\sum_{1\leqq k<n\leqq r}(kn)^{-\sigma}\left|\int_{t'}^{t} u^{\sigma-\frac{1}{2}}\left(\frac{n}{k}\right)^{iu}du\right|\right),$$

where $t' = \max\left(t_1, \frac{2\pi n^2}{m}\right)$, $r = r(t)$. Moreover

$$\int_{t'}^{t} u^{\sigma-\frac{1}{2}}\left(\frac{n}{k}\right)^{iu}du = \frac{1}{\log\frac{n}{k}}\,O(t^{\sigma-\frac{1}{2}})$$

and

$$\sum_{1\leqq k<n\leqq r}(kn)^{-\sigma}/\log\frac{n}{k} \leqq \frac{1}{\log 2}\sum_{k\leqq\frac{n}{2}}(kn)^{-\sigma} + \sum_{\frac{n}{2}<k<n}(kn)^{-\sigma}\frac{n}{n-k}$$
$$= O(1)\sum_{n=1}^{r} n^{1-2\sigma} + O(1)\sum_{n=1}^{r} n^{1-2\sigma}\log n = O(t^{1-\sigma}\log t).$$

Hence

$$\Delta^{-2}H^2 < \frac{\varphi(m)}{m(1-2\sigma)}(1+o(1)) + \Delta^{-1}O(t^{\frac{1}{2}}\log t) = \frac{\varphi(m)}{m(1-2\sigma)}(1+o(1)),$$

and the assertion follows from (50), (51).

Consider now the remaining case $t \geqq 2t_1$. We set $\log\frac{t}{t_1}\Big/\log 2 = h, [h]+1 = h_0$ and apply (49) for the h_0 intervals $u_k \leqq u \leqq u_{k+1}$, $u_k = t_1 2^{\frac{kl}{h_0}}$ $(k = 0, \cdots, h_0-1)$; the assertion follows by summation over k.

THEOREM 7: *If* $t^{\frac{1}{2}}\log t = o(t-t_1)$, *then* $\liminf_{t\to\infty} A(t_1, t)/(t-t_1) \geqq m/\pi e\varphi(m)$.

PROOF: Apply Lemmata 14, 15 with $\sigma_0 = \sigma = \frac{1}{2}\left(1 - \frac{e\varphi(m)}{m}\right)$; then

$$\frac{\pi e\varphi(m)}{2m} A(t_1, t) > -\frac{t - t_1}{2} (\log e^{-1} + o(1)) = \frac{t - t_1}{2} (1 + o(1));$$

q.e.d.

We denote by $B(t_1, t_2, \epsilon)$ the number of zeros of $L(s)$ in the rectangle $t_1 < t < t_2$, $\frac{1}{2} - \epsilon < \sigma < \frac{1}{2} + \epsilon$ ($\epsilon > 0$, $0 < t_1 < t_2$).

THEOREM 8: *If* $t^{\frac{1}{2}} \log t = o(t - t_1)$ *and* $\epsilon = o(\log\log t/\log t)$, *then*

$$\liminf_{t\to\infty} B(t_1, t, \epsilon)/(t - t_1)t^{\epsilon} \geqq m/4\pi e\varphi(m).$$

PROOF: Let $B^*(t_1, t_2, \epsilon)$ be the number of zeros of $L(s)$ with $\sigma \geqq \frac{1}{2} + \epsilon$, $t_1 < t < t_2$, and put $B(t_1, t, \epsilon) = B$, $B^*(t_1, t, \epsilon) = B^*$, $A(t_1, t) = A$. We have

$$B \geqq A \tag{52}$$

and

$$B + 2B^* = \frac{1}{2\pi}\int_{t_1}^{t} \log\frac{mu}{2\pi}\, du + O(\log t) > \frac{\Delta}{2\pi}\log\frac{mt}{2\pi e} + O(\log t). \tag{53}$$

On the other hand, by (45),

$$|\beta(\tfrac{1}{2} + ti)| \geqq |\Re\{\beta(\tfrac{1}{2} + ti)\}| = \tfrac{1}{2}|L(\tfrac{1}{2} + ti)|,$$

by (46),

$$\int_{t_1}^{t} \log|L(\tfrac{1}{2} + ui)|\, du > 2\pi\epsilon B^* + O(\log t).$$

Applying (47), (48) and Lemma 15, we obtain

$$\pi(\tfrac{1}{2} - \sigma)A > 2\pi\epsilon B^* - \Delta\log 2 - \frac{\Delta}{2}\left(\log\frac{\varphi(m)}{m(1 - 2\sigma)} + o(1)\right),$$

where $\sigma_1 \leqq \sigma < \frac{1}{2}$ and $\Delta = t - t_1$. Set $4\pi\epsilon B^*\Delta^{-1} = \eta$ and choose

$$\sigma = \tfrac{1}{2} - \frac{2\varphi(m)}{\gamma.t} e^{\gamma-\eta}:$$

then

$$A > \frac{m}{4\pi\varphi(m)} \Delta e^{\eta-1}(1 + o(1)).$$

By (52), (53),

$$2\pi\Delta^{-1}B > \max\left\{\frac{m}{2\varphi(m)} e^{\eta-1}(1 + o(1)), \log\frac{mt}{2\pi e} - \epsilon^{-1}\eta\right\} + \Delta^{-1}O(\log t).$$

Putting

$$\eta_0 = \epsilon\log\frac{mt}{2\pi e} - \frac{m\epsilon t^{\epsilon}}{2e\varphi(m)},$$

we infer that

$$\log \frac{mt}{2\pi e} - \epsilon^{-1}\eta \geqq \log \frac{mt}{2\pi e} - \epsilon^{-1}\eta_0 = \frac{mt^\epsilon}{2e\varphi(m)} \qquad (\eta \leqq \eta_0),$$

$$e^\eta > e^{\eta_0} = \left(\frac{mt}{2\pi e}\right)^\epsilon (1 + o(1)) = t^\epsilon(1 + o(1)) \qquad (\eta > \eta_0),$$

whence

$$2\pi\Delta^{-1}B > \frac{mt^\epsilon}{2e\varphi(m)} (1 + o(1));$$

q.e.d.

Part II: Epstein zeta-functions

6. Modular forms and Dirichlet series

Let $\Phi(z)$ be a modular form of weight g, with the multiplier system $v = v(a, b, c, d)$ for a subgroup Δ of the modular group Γ, of finite index; this means

$$(cz + d)^{-g}\Phi\left(\frac{az + b}{cz + d}\right) = v\Phi(z) \tag{54}$$

for all substitutions in Δ, and

$$(cz + d)^{-g}\Phi\left(\frac{az + b}{cz + d}\right) = \sum_{n=0}^{\infty} a_n e^{\frac{2\pi i n}{q} z} \qquad (\Im\{z\} > 0) \tag{55}$$

for all substitutions in Γ, where the coefficients $a_0, a_1, \cdots$ depend upon a, b, c, d and q is a uniquely determined positive integer. The modular form $\Phi(z)$ is called a cusp-form, if $a_0 = 0$ for all modular substitutions.

With any modular form $\Phi(z)$ we may associate a Dirichlet series $\hat{\Phi}(s)$ in the following way: We consider the expansion (55) in the case of the identical substitution,

$$\Phi(z) = \sum_{n=0}^{\infty} a_n e^{\frac{2\pi i n}{q} z}, \tag{56}$$

and define

$$\hat{\Phi}(s) = \sum_{n=1}^{\infty} a_n n^{-s}. \tag{57}$$

Lemma 16: *Let $\Phi(z)$ be a modular cusp-form of weight g, and let all its multipliers have the absolute value* 1; *then $\hat{\Phi}(s)$ is an entire function and $\hat{\Phi}(s) = t^{\frac{1}{2}} \log t\, O(1 + t^{g-2\sigma})$.*

Proof: It is known that $a_k = O\left(k^{\frac{g}{2}}\right)$, hence (57) converges at least in the half-plane $\sigma > \frac{g}{2} + 1$. Let $t > 2$, $\alpha = e^{\pi i(1-t^{-1})}$, $\left(\frac{iq}{2\pi\alpha}\right)^s \Gamma(s)\hat{\Phi}(s) = \varphi(s)$; then

$$\varphi(s) = \int_0^\infty r^{s-1}\Phi(\alpha r)\,dr. \tag{58}$$

If y denotes the imaginary part of z, then $y^{\frac{g}{2}}\,|\Phi(z)|$ is invariant under the substitutions of Δ, by (54), and vanishes in the parabolic frontier-points of a fundamental domain of Δ, by (55), whence

$$|\Phi(z)| < c_{19}\,y^{-\frac{g}{2}}.$$

Now $\Im\{\alpha r\} = r\sin\dfrac{\pi}{t} > \dfrac{2r}{t}$, and consequently

$$|\Phi(\alpha r)| < c_{19}\left(\frac{t}{2r}\right)^{\frac{g}{2}}. \tag{59}$$

On the other hand, by (56) and the corresponding formula (55) for the substitution $z \to -z^{-1}$, we obtain, since $\Im\{-\alpha^{-1}r^{-1}\} = r^{-1}\sin\dfrac{\pi}{t} > \dfrac{2}{rt}$,

$$|\Phi(\alpha r)| < c_{20}e^{-\frac{4\pi r}{qt}} \quad (r \geqq t), \qquad |\Phi(\alpha r)| < c_{21}r^{-g}e^{-\frac{4\pi}{qrt}} \quad (r \leqq t^{-1}). \tag{60}$$

We use (59), for $t^{-1} \leqq r \leqq t$, and (60). It follows that $\varphi(s)$ is regular in the whole s-plane, by (58), and that

$$\varphi(s) = O\left(\int_0^{t^{-1}} r^{\sigma-g-1}e^{-\frac{4\pi}{qrt}}\,dr + t^{\frac{g}{2}}\int_{t^{-1}}^{t} r^{\sigma-\frac{g}{2}-1}\,dr + \int_t^\infty r^{\sigma-1}e^{-\frac{4\pi r}{qt}}\,dr\right)$$

$$= O\left(t^{g-\sigma} + t^{\frac{g}{2}}\left(t^{\sigma-\frac{g}{2}} + t^{\frac{g}{2}-\sigma}\right)\log t + t^\sigma\right) = t^\sigma \log t\, O(1 + t^{g-2\sigma}).$$

Since $\left(\dfrac{2\pi\alpha}{iq}\right)^s \Big/ \Gamma(s) = O(t^{\frac{1}{2}-\sigma})$, the assertion is proved.

7. Application of the analytic theory of quadratic forms

Let $\mathfrak{S}$ be an integral positive symmetric matrix with k rows and denote by $\alpha(n)$ the number of integral solutions $\mathfrak{x}$ of $\mathfrak{x}'\mathfrak{S}\mathfrak{x} = n$, where n is any integer. It is known that $f(z;\mathfrak{S}) = f(z) = \sum_{n=0}^{\infty}\alpha(n)e^{\pi i n z}$ is a modular form of weight $\dfrac{k}{2}$ whose multipliers are roots of unity depending upon the genus γ of $\mathfrak{S}$; moreover, if also $\mathfrak{S}_1$ belongs to γ, then the difference $f(z;\mathfrak{S}) - f(z;\mathfrak{S}_1)$ is a cusp-form. The corresponding Dirichlet series $\hat{f}(s) = \sum_{n=1}^{\infty}\alpha(n)n^{-s} = \zeta(s;\mathfrak{S})$ are Epstein zeta-functions. On account of Lemma 16, $\zeta(s;\mathfrak{S}) - \zeta(s;\mathfrak{S}_1) = t^{\frac{1}{2}}\log t\, O\left(1 + t^{\frac{k}{2}-2\sigma}\right)$, whenever $\mathfrak{S}$, $\mathfrak{S}_1$ are in the same genus.

Let $\mathfrak{S}_1, \cdots, \mathfrak{S}_h$ denote representatives of all different classes of γ and define

$$F(z;\gamma) = F(z) = \sum_{l=1}^{h}\frac{f(z;\mathfrak{S}_l)}{E(\mathfrak{S}_l)} \Big/ \sum_{l=1}^{h}\frac{1}{E(\mathfrak{S}_l)}, \qquad \hat{F}(s) = \mathrm{Z}(s;\gamma),$$

where $E(\mathfrak{S})$ denotes the number of units of $\mathfrak{S}$. Since

$$\zeta(s;\mathfrak{S}) - \mathrm{Z}(s;\gamma) = \sum_{l=1}^{h} \frac{\zeta(s;\mathfrak{S}) - \zeta(s;\mathfrak{S}_l)}{E(\mathfrak{S}_l)} \Big/ \sum_{l=1}^{h} \frac{1}{E(\mathfrak{S}_l)},$$

we infer

THEOREM 9:

$$\zeta(s;\mathfrak{S}) = \mathrm{Z}(s;\gamma) + t^{\frac{1}{2}} \log t \, O\Big(1 + t^{\frac{k}{2}-2\sigma}\Big).$$

The matrix $\mathfrak{S}$ is called even, if $\alpha(n) = 0$ for all odd integers n. Let D be the determinant of $\mathfrak{S}$ and define, for any pair of coprime integers a, b and $b > 0$,

$$H\left(\frac{a}{b};\gamma\right) = H\left(\frac{a}{b}\right) = D^{-\frac{1}{2}} b^{-\frac{k}{2}} \sum_{r \bmod b} e^{\pi i \frac{a}{b} t' \mathfrak{S} r},$$

if $ab\,\mathfrak{S}$ is even, and $H\left(\frac{a}{b}\right) = 0$ otherwise.

THEOREM 10: *If* $1 < \sigma < \frac{k}{2} - 1$, *then*

$$\mathrm{Z}(s;\gamma) = \pi^s \frac{\Gamma\left(\frac{k}{2} - s\right)}{\Gamma\left(\frac{k}{2}\right)} \sum_{a,b} a^{s-\frac{k}{2}} b^{-s} \left\{ e^{\frac{\pi i}{4}(2s-k)} H\left(\frac{a}{b}\right) + e^{\frac{\pi i}{4}(k-2s)} H\left(\frac{-a}{b}\right) \right\},$$

where a, b run over all pairs of positive coprime integers.

PROOF: Put $e^{\frac{\pi i}{2}s} \pi^{-s} \Gamma(s)\mathrm{Z}(s;\gamma) = g(s)$; then

$$g(s) = \int_0^\infty z^{s-1}(F(z) - 1)\, dz \qquad \left(\sigma > \frac{k}{2}\right),$$

where the integration is extended over the positive imaginary axis $z = iy$.

It is known that

$$F(z) - 1 = e^{\frac{\pi i k}{4}} \sum_{a,b} H\left(\frac{a}{b}\right)(bz - a)^{-\frac{k}{2}} \qquad (k < 4), \tag{61}$$

where a, b run over all pairs of coprime integers with $b > 0$, and that $\left|H\left(\frac{a}{b}\right)\right| \leqq 1$. Since the sum $\sum_{a,b} (a^2 + b^2)^{-\frac{k}{4}} \max\left(1, y^{-\frac{k}{2}}\right)$ is a dominant series for the expansion (61) with $z = iy$, we obtain

$$g(s) = \frac{H(1) e^{\frac{\pi i}{2}s} Y^{s-\frac{k}{2}}}{s - \frac{k}{2}} + e^{\frac{\pi i k}{4}} \sum_{\substack{a,b \\ a \neq 0}} H\left(\frac{a}{b}\right) \int_0^{iY} z^{s-1}(bz - a)^{-\frac{k}{2}}\, dz$$

$$+ \int_{iY}^\infty z^{s-1}(F(z) - 1)\, dz \qquad \left(\sigma > \frac{k}{2}\right), \tag{62}$$

for any $Y > 0$. The last term in (62) is an entire function of s. Now

$$\sum_{\substack{a,b \\ a\neq 0}} \int_0^Y y^{\sigma-1}(b^2y^2 + a^2)^{-\frac{k}{4}}dy < \sum_{a\neq 0} \int_0^\infty \left(\int_0^Y y^{\sigma-1}(b^2y^2 + a^2)^{-\frac{k}{4}}dy\right)db$$

$$= 2\zeta\left(\frac{k}{2} - 1\right)\frac{Y^{\sigma-1}}{\sigma - 1}\int_0^\infty (b^2 + 1)^{-\frac{k}{4}}db, \tag{63}$$

for $\sigma > 1$; therefore the infinite series in (62) converges uniformly for $1 < \sigma_1 \leqq \sigma \leqq \sigma_2$ and any fixed $Y > 0$. This proves that (62) holds good for $\sigma > 1$. Assume $1 < \sigma < \frac{k}{2} - 1$ and use (63) with Y^{-1}, $\frac{k}{2} - \sigma$ instead of Y, σ; it follows, for $Y = 1$, that

$$\sum_{\substack{a,b \\ a\neq 0}} \int_0^\infty y^{\sigma-1}(b^2y^2 + a^2)^{-\frac{k}{4}}dy$$

$$< 2\zeta\left(\frac{k}{2} - 1\right)\left(\frac{1}{\sigma - 1} + \frac{1}{\frac{k}{2} - \sigma - 1}\right)\int_0^\infty (b^2 + 1)^{-\frac{k}{4}}db,$$

and consequently the series in (62) converges also uniformly with respect to Y. Performing the passage to the limit $Y \to \infty$, we obtain

$$g(s) = e^{\frac{\pi i k}{4}} \sum_{a,b} \left\{H\left(\frac{a}{b}\right)\int_0^\infty z^{s-1}(bz - a)^{-\frac{k}{2}}dz \right.$$

$$\left. + H\left(\frac{-a}{b}\right)\int_0^\infty z^{s-1}(bz + a)^{-\frac{k}{2}}dz\right\} \quad \left(1 < \sigma < \frac{k}{2} - 1\right),$$

where a, b run over all pairs of positive coprime integers. Since

$$e^{\pi i\left(\frac{k}{2}-s\right)}\int_0^\infty z^{s-1}(bz - a)^{-\frac{k}{2}}dz = \int_0^\infty z^{s-1}(bz + a)^{-\frac{k}{2}}dz = a^{s-\frac{k}{2}}b^{-s}\frac{\Gamma(s)\Gamma\left(\frac{k}{2} - s\right)}{\Gamma\left(\frac{k}{2}\right)},$$

the assertion follows.

Henceforth we assume $D = 1$. If k is not divisible by 8, then there exists exactly one genus, the genus γ_k of the unit matrix $\mathfrak{E}_k$. If k is divisible by 8, then there exist exactly two genera, namely γ_k and a genus γ_k^* of even matrices.

THEOREM 11:

$$Z(s; \gamma_k^*) = \frac{(2\pi)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)\zeta\left(\frac{k}{2}\right)}2^{-s}\zeta(s)\zeta\left(s + 1 - \frac{k}{2}\right).$$

PROOF: It follows from the known values of the Gaussian sums that $H\left(\frac{a}{b}; \gamma_k^*\right) = 1$. By Theorem 10,

$$Z(s;\gamma_k^*) = 2\pi^s \frac{\Gamma\left(\frac{k}{2}-s\right)}{\Gamma\left(\frac{k}{2}\right)} \cos\frac{\pi s}{2}\,\zeta(s)\,\zeta\left(\frac{k}{2}-s\right)\Big/\zeta\left(\frac{k}{2}\right),$$

and the assertion follows from the functional equation of $\zeta(s)$.

THEOREM 12:

$$Z(s;\gamma_k) = 2\pi^s \frac{\Gamma\left(\frac{k}{2}-s\right)}{\Gamma\left(\frac{k}{2}\right)}\left\{\psi(s)+\psi\left(\frac{k}{2}-s\right)\right\} \qquad \left(1<\sigma<\frac{k}{2}-1\right), \tag{64}$$

where

$$\psi(s) = 2^{s-\frac{k}{2}}\left\{\cos\frac{\pi}{4}(2s-k)\sum_{\substack{a,b\\ b\equiv 1 \pmod 4}} \chi_b^k(a)a^{s-\frac{k}{2}}b^{-s} + \cos\frac{\pi}{4}(2s+k)\sum_{\substack{a,b\\ b\equiv 3 \pmod 4}} \chi_b^k(a)a^{s-\frac{k}{2}}b^{-s}\right\},$$

$\chi_b(a) = \left(\frac{a}{b}\right)$ *denoting the Legendre-Jacobi symbol.*

PROOF: We have $H\left(\frac{a}{b};\gamma_k\right) = 0$, if ab is odd,

$$H\left(\frac{a}{b}\right) = i^{k\left(\frac{b-1}{2}\right)^2}\chi_b^k(2a) \qquad (a \text{ even}, b \text{ odd}), \tag{65}$$

$$H\left(\frac{a}{b}\right) = e^{\frac{\pi i k}{4}} H\left(\frac{-b}{a}\right), \quad H\left(\frac{-a}{b}\right) = e^{-\frac{\pi i k}{4}} H\left(\frac{b}{a}\right) \quad (0<a \text{ odd}, b \text{ even}). \tag{66}$$

Define

$$q(s) = \sum_{\substack{a,b\\ a \text{ even}}} a^{s-\frac{k}{2}}b^{-s}\left\{e^{\frac{\pi i}{4}(2s-k)}H\left(\frac{a}{b}\right)+e^{\frac{\pi i}{4}(k-2s)}H\left(\frac{-a}{b}\right)\right\},$$

then, by Theorem 10 and (66),

$$Z(s;\gamma_k) = \pi^s \frac{\Gamma\left(\frac{k}{2}-s\right)}{\Gamma\left(\frac{k}{2}\right)}\left\{q(s)+q\left(\frac{k}{2}-s\right)\right\}.$$

On the other hand, by (65),

$$q(s) = 2^{s-\frac{k}{2}}\sum_{\substack{a,b\\ b \text{ odd}}} a^{s-\frac{k}{2}}b^{-s}i^{k\left(\frac{b-1}{2}\right)^2}\chi_b^k(a)\left(e^{\frac{\pi i}{4}(2s-k)}+\chi_b^k(-1)e^{\frac{\pi i}{4}(k-2s)}\right) = 2\psi(s);$$

q.e.d.

If k is even, then (64) may be expressed in a different way, analogous to Theorem 11. Obviously

$$\psi(s) = 2^{s-\frac{k}{2}} \cos\frac{\pi}{4}(2s-k) \sum_{\substack{a,b \\ b \text{ odd}}} (-1)^{\frac{k(b-1)}{4}} a^{s-\frac{k}{2}} b^{-s} \qquad (k \text{ even}),$$

whence

$$\psi(s) = 2^{s-\frac{k}{2}} \cos\frac{\pi}{4}(2s-k)\,\zeta\left(\frac{k}{2}-s\right)\zeta(s)(1-2^{-s})\Big/\zeta\left(\frac{k}{2}\right)\left(1-2^{-\frac{k}{2}}\right)$$

$$(k \equiv 0 \pmod 4),$$

$$\psi(s) = 2^{s-\frac{k}{2}} \cos\frac{\pi}{4}(2s-k)\,\zeta\left(\frac{k}{2}-s\right) L_{-4}(s)\Big/L_{-4}\left(\frac{k}{2}\right) \qquad (k \equiv 2 \pmod 4),$$

with $L_{-4}(s) = \sum_{n=0}^{\infty} (-1)^n (2n+1)^{-s}$. By (64) and the functional equations of $\zeta(s)$, $L_{-4}(s)$, we obtain

$$\mathrm{Z}(s;\gamma_k) = \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)\zeta\left(\frac{k}{2}\right)\left(1-2^{-\frac{k}{2}}\right)}$$

$$\cdot\left\{1-2^{-s}+(-1)^{\frac{k}{4}}\left(2^{\frac{k}{2}-2s}-2^{-s}\right)\right\}\zeta(s)\,\zeta\left(s+1-\frac{k}{2}\right) \qquad (k \equiv 0),$$

$$\mathrm{Z}(s;\gamma_k) = \frac{\pi^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right) L_{-4}\left(\frac{k}{2}\right)}$$

$$\cdot\left\{L_{-4}(s)\,\zeta\left(s+1-\frac{k}{2}\right)+(-1)^{\frac{k-2}{4}}\,2^{1-\frac{k}{2}}\,\zeta(s) L_{-4}\left(s+1-\frac{k}{2}\right)\right\} \qquad (k \equiv 2).$$

For odd values of k, the corresponding formula is more complicated. Let d be a discriminant, i.e. an integer such that either $d \equiv 1 \pmod 4$ and d not divisible by any square $\neq 1$, or $d \equiv 8, 12 \pmod 4$ and $\frac{d}{4}$ not divisible by any square $\neq 1$, and define $L_d(s) = \sum_{n=1}^{\infty} \left(\frac{d}{n}\right) n^{-s}$; then it follows from Theorem 12 and the functional equation of $L_d(s)$ that

$$\pi^{-\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)\mathrm{Z}(s;\gamma_k)/\zeta(2s)$$

$$= (1-2^{-2s}) \sum_{d \equiv 1 \pmod 4} d^{\frac{1-k}{2}} L_d\left(s+1-\frac{k}{2}\right)\Big/ L_d\left(s+\frac{k}{2}\right)\left(1-\left(\frac{d}{2}\right)2^{-s-\frac{k}{2}}\right)$$

$$+\cos\frac{\pi k}{4}\,2^{\frac{k}{2}-2s}\sum_{d\equiv 1 \pmod 4} d^{\frac{1-k}{2}}\,L_d\left(s+1-\frac{k}{2}\right)\left(1-\left(\frac{d}{2}\right)2^{s-\frac{k}{2}}\right)\Big/L_d\left(s+\frac{k}{2}\right)$$

$$+\cos\frac{\pi k}{4}\,2^{\frac{k}{2}}\sum_{d\equiv 0 \pmod 4} d^{\frac{1-k}{2}}\,L_d\left(s+1-\frac{k}{2}\right)\Big/L_d\left(s+\frac{k}{2}\right)\qquad\left(\sigma>\frac{k}{2}\right),$$

where the summation extends over all discriminants d. This formula is of some interest, because it connects the L-series with the theory of the Epstein zeta-functions, but we do not need it for our further purpose and omit the detailed proof.

8. The zeros of $\zeta(s;\mathfrak{S})$ in the strip $1<\sigma<\frac{k}{2}-1$.

THEOREM 13: *Let $\mathfrak{S}$ be in the genus γ_k^* and $0<\epsilon<\frac{k}{4}-1$; then the number of zeros of $\zeta(s;\mathfrak{S})$ in the strip $1+\epsilon\leqq\sigma\leqq\frac{k}{2}-1-\epsilon$ is finite.*

PROOF: By Theorems 9 and 11, we have

$$\zeta(s;\mathfrak{S})=\mathrm{Z}(s;\gamma_k^*)\left\{1+t\log t\,O\left(t^{\sigma-\frac{k}{2}}+t^{-\sigma}\right)\right\}$$

$$=\mathrm{Z}(s;\gamma_k^*)\{1+\log t\,O(t^{-\epsilon})\}\qquad\left(1+\epsilon\leqq\sigma\leqq\frac{k}{2}-1-\epsilon\right),$$

and $\mathrm{Z}(s;\gamma_k^*)$ has exactly $\frac{k}{4}-2$ zeros in the strip $1\leqq\sigma\leqq\frac{k}{2}-1$, namely $s=3,5,\cdots,\frac{k}{2}-3$.

THEOREM 14: *Let $\mathfrak{S}$ be in the genus γ_k and $k\geqq 12$; then all zeros of $\zeta(s;\mathfrak{S})$ in the strip $2\leqq\sigma\leqq\frac{k}{2}-2$ are simple and lie on $\sigma=\frac{k}{4}$, with at most a finite number of exceptions, and the interval $0<u<t$ contains exactly $\frac{t}{\pi}\log 2+O(1)$ zeros of $\zeta\left(\frac{k}{4}+ui;\mathfrak{S}\right)$.*

PROOF: By Theorems 9, 12,

$$\zeta(s;\mathfrak{S})=2\pi^s\frac{\Gamma\left(\frac{k}{2}-s\right)}{\Gamma\left(\frac{k}{2}\right)}\left\{\psi(s)+\psi\left(\frac{k}{2}-s\right)\right\}$$

$$+t^{\frac{1}{2}}\log t\,O\left(1+t^{\frac{k}{2}-2\sigma}\right)\qquad\left(1<\sigma<\frac{k}{2}-1\right),\tag{67}$$

$$\zeta(s;\mathfrak{S})=\pi^s\frac{\Gamma\left(\frac{k}{2}-s\right)}{\Gamma\left(\frac{k}{2}\right)}\,e^{\frac{\pi i}{8}(k-4s)}\tag{68}$$

$$\left\{\rho(s) + \overline{\rho\left(\frac{k}{2} - \bar{s}\right)} + \log t\, O(t^{-1})\right\} \quad \left(2 \leqq \sigma \leqq \frac{k}{2} - 2\right),$$

where

$$\rho(s) = e^{\frac{\pi i k}{8}} 2^{s-\frac{k}{2}} \left\{ \sum_{\substack{a,b \\ b \equiv 1 \pmod 4}} \chi_b^k(a) a^{s-\frac{k}{2}} b^{-s} - i^k \sum_{\substack{a,b \\ b \equiv 3 \pmod 4}} \chi_b^k(a) a^{s-\frac{k}{2}} b^{-s} \right\}.$$

Put $e^{-\frac{\pi i k}{8}} 2^{\frac{k}{2}-s} \rho(s) = 1 + R$, then

$$1 + |R| \leqq \sum_{\substack{a,b \\ b \text{ odd}}} a^{\sigma-\frac{k}{2}} b^{-\sigma} = \zeta\left(\frac{k}{2} - \sigma\right) \zeta(\sigma)(1 - 2^{-\sigma}) / \zeta\left(\frac{k}{2}\right)\left(1 - 2^{-\frac{k}{2}}\right); \tag{69}$$

and consequently

$$\rho(s) + \overline{\rho\left(\frac{k}{2} - \bar{s}\right)} = \begin{cases} 2^{s-\frac{k}{2}} e^{\frac{\pi i k}{8}} (1 + R_1) & \left(\sigma = \frac{k}{2} - 2\right) \\ 2^{-s} e^{-\frac{\pi i k}{8}} (1 + R_1) & (\sigma = 2) \end{cases} \tag{70}$$

with

$$|R_1| < -1 + \frac{\zeta(2)\zeta\left(\frac{k}{2} - 2\right)\left(1 - 2^{2-\frac{k}{2}}\right)}{\zeta\left(\frac{k}{2}\right)\left(1 - 2^{-\frac{k}{2}}\right)}$$

$$+ 2^{4-\frac{k}{2}} \left(-1 + \frac{\zeta\left(\frac{k}{2} - 2\right)\zeta(2)(1 - 2^{-2})}{\zeta\left(\frac{k}{2}\right)\left(1 - 2^{-\frac{k}{2}}\right)} \right) \tag{71}$$

$$\leqq \frac{\zeta(2)\zeta(4)}{\zeta(6)(1 - 2^{-6})}(1 + 2^{-3}) - (1 + 2^{-2}) = \frac{3}{4} \qquad (k \geqq 12).$$

Let $\Gamma\left(\frac{k}{2}\right)\pi^{-s} e^{\frac{\pi i}{8}(4s-k)} \zeta(s; \mathfrak{S}) \Big/ \Gamma\left(\frac{k}{2} - s\right) = g(s)$; by (68), (70), (71), the function $\log g(s) - \left(s - \frac{k}{2}\right) \log 2$ is regular and bounded on $\sigma = \frac{k}{2} - 2$ for sufficiently large values of t, and the same holds for $\log g(s) + s \log 2$ on $\sigma = 2$. Analogous to Lemma 13, it is easily proved, by (68), that the variation of $\log g(s)$ on the segment $s = \sigma + ti$ is bounded, if σ runs from 2 to $\frac{k}{2} - 2$. Let $B(t_1)$ be the number of zeros of $\zeta(s; \mathfrak{S})$ in the rectangle $0 < t < t_1$, $2 \leqq \sigma \leqq \frac{k}{2} - 2$; then $2\pi i\, B(t_1)$ is the variation of $\log g(s)$ on the contour, whence

$$B(t) = \frac{t}{\pi} \log 2 + O(1) \qquad (k \geqq 12). \tag{72}$$

On the other hand, by the functional equation of $\zeta(s; \mathfrak{S})$, we have

$$\zeta(s; \mathfrak{S}) = 2\pi^s \frac{\Gamma\left(\frac{k}{2} - s\right)}{\Gamma\left(\frac{k}{2}\right)} e^{\frac{\pi i}{8}(k-4s)} \Re\{h(s)\} \qquad \left(\sigma = \frac{k}{4}\right),$$

where $h(s)$ is regular on $\sigma = \frac{k}{4}$ and, by (67),

$$h(s) = \rho(s) + \log t\, O(t^{1-\frac{k}{4}}) \qquad (k > 4).$$

By (69),

$$\left| e^{-\frac{\pi i k}{8}} 2^{\frac{k}{2}-s} \rho(s) - 1 \right| \leqq -1 + \zeta^2\left(\frac{k}{4}\right) \Big/ \zeta\left(\frac{k}{2}\right)\left(1 + 2^{-\frac{k}{4}}\right)$$

$$< -1 + \zeta^2(2)/\zeta(4)(1 + 2^{-2}) = 1 \qquad (k > 8);$$

hence

$$\log h(s) = ti \log 2 + O(1) \qquad \left(\sigma = \frac{k}{4}, k > 8\right).$$

Let $A(t_1)$ denote the number of different zeros of odd order of $\zeta(s; \mathfrak{S})$ in the interval $0 < t < t_1$ on the line $s = \frac{k}{4} + ti$. By Lemma 12,

$$A(t) > \frac{t}{\pi} \log 2 + O(1) \qquad (k > 8),$$

and the assertion follows from (72).

THE INSTITUTE FOR ADVANCED STUDY

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

ON CONTINUOUS PATH-SURFACES OF ZERO AREA

BY TIBOR RADÓ

(Received December 28, 1942)

INTRODUCTION

0.1. If a curve C, in terms of Cartesian coordinates x_1, x_2, x_3, is given by equations of the form

$$x_i = x_i(t), \qquad i = 1, 2, 3, \qquad t' \leqq t \leqq t'', \tag{1}$$

then the points (x_1, x_2, x_3), obtained by means of these equations, form a certain point-set E. However, the length of the curve cannot be deduced from a mere knowledge of E. For example, we may assign E as a straight segment which is described, for suitable choice of the equations (1), several times in either direction while t varies from t' to t''. Hence it is clear that the length of a *path-curve*, that is a curve determined by equations of the form (1), is determined by the manner in which the point-set E is described, rather than by the point-set E itself. The purpose of the term *path-curve* is precisely to call attention to this fact.

0.2. Similar considerations apply to surfaces. A *continuous path-surface* S is determined, in terms of Cartesian coordinates x_1, x_2, x_3, by equations of the form

$$S: \qquad x_i = x_i(P), \qquad i = 1, 2, 3, \qquad P \in U, \tag{2}$$

where $x_1(P)$, $x_2(P)$, $x_3(P)$ are continuous functions of the point P on the surface U of the unit sphere. The term path-surface has been proposed by C. B. Morrey (see Morrey [1] in the Bibliography at the end of this paper). In the literature on the area of surfaces, the range of definition of the coordinate functions is usually taken as a plane Jordan region, rather than the surface U of the unit sphere. The discussion is quite analogous in either case, but it is more convenient for our present purposes to work on the unit sphere. It will be also convenient to use vector notation. Instead of (2), we shall write

$$S: \qquad \mathfrak{x} = \mathfrak{x}(P), \qquad P \in U, \tag{3}$$

where $\mathfrak{x}$ stands for the vector with components x_1, x_2, x_3.

0.3. Suppose there exists a triangulation T of U, comprised of (curvilinear) triangles $\Delta_1, \Delta_2, \cdots, \Delta_m$, such that each Δ_j is carried, by means of (3), in a biunique way into a plane rectilinear triangle Δ_j' in (x_1, x_2, x_3) space. In this case, we shall call S a *polyhedral path-surface*, and we shall designate by $E(S)$ the area of S in the elementary sense, that is, the sum of the areas of the triangles Δ_j'.

0.4. It will be convenient to consider also another type of polyhedral surfaces, to be termed simply *polyhedra*, in contradistinction to the polyhedral path-surfaces defined in 0.3. Let again T be a triangulation of U, comprised of the

(curvilinear) triangles $\delta_1, \delta_2, \cdots, \delta_n$. With each vertex V of the triangulation T let there be associated a point V' in (x_1, x_2, x_3) space. With each triangle δ_j of T we associate the rectilinear triangle δ_j', in (x_1, x_2, x_3) space, whose vertices are the three points V' which were associated with the three vertices V of δ_j. The triangles δ_j' constitute then a *polyhedron* $\mathfrak{P}$ and the sum of the areas of $\delta_1', \delta_2', \cdots, \delta_n'$ is the area $E(\mathfrak{P})$ of $\mathfrak{P}$. The usefulness of these definitions will become apparent in the sequel. It should be observed that we do not require that the correspondence between the vertices V and V' be bi-unique. As a consequence, some or all of the triangles δ_j' may reduce to straight segments or even to single points. If this happens for a certain triangle δ_j', then its area is taken as equal to zero.

0.5. The definition of the Lebesgue area $A(S)$ of the continuous path-surface (3) may now be stated in the following two forms which are easily shown to be equivalent. Let

$$S_n: \qquad \mathfrak{x} = \mathfrak{x}_n(P), \qquad P \in U,$$

be a sequence of polyhedral path-surfaces such that

$$\mathfrak{x}_n(P) \to \mathfrak{x}(P)$$

uniformly on U. Then

$$A(S) = \text{gr. l. b. } \underline{\lim}\, E(S_n), \tag{4}$$

where the greatest lower bound is taken with respect to all sequences S_n with the properties just stated. The reader interested in learning about recent literature on the Lebesgue area $A(S)$, may consult Radó [2].

0.6. Given a continuous path-surface S as in (3), and a polyhedron $\mathfrak{P}$, let us introduce a quantity $d(S, \mathfrak{P})$ which will in a sense measure the deviation of $\mathfrak{P}$ from S, as follows. According to 0.4, $\mathfrak{P}$ arises from a triangulation T of U, comprised of certain curvilinear triangles $\delta_1, \delta_2, \cdots, \delta_n$. For each vertex V of T, we have a corresponding vertex V' of $\mathfrak{P}$. On the other hand, to each vertex V of T there corresponds a point $\bar{V}$ by means of (3). We define $d(S, \mathfrak{P})$ as the maximum distance between the points V' and $\bar{V}$ for all vertices V of T, *plus* the maximum diameter of the triangles $\delta_1, \delta_2, \cdots, \delta_n$ of T. We have then the formula, easily recognized as equivalent to (4),

$$A(S) = \text{gr. l. b. } \underline{\lim}\, E(\mathfrak{P}_n),$$

where the greatest lower bound is taken with respect to all sequences of polyhedra $\mathfrak{P}_n$ such that $d(S, \mathfrak{P}_n) \to 0$.

0.7. It is obvious that a continuous path-curve has zero length if and only if it reduces to a single point. The analogous question for continuous path-surfaces, to be referred to for brevity as *the characterization problem*, is of considerable complexity, and also of considerable importance as an apparently indispensable step in approaching various unsolved fundamental problems in the theory of the area of surfaces. The characterization problem has been

studied extensively by Geöcze (see Geöcze) [1], [2], [3], [4]). His main contributions are described in sections 3.16, 2.9, 2.14, 1.21, 3.5 of this paper. The recent work of Morrey on generalized conformal maps of surfaces (Morrey [1]) implies several interesting results concerning our topic. For example, the theorems in 2.8 and 2.13 could be also deduced from the general results of Morrey. Since his work depends to a large extent upon conformal mapping and Lebesgue theory, it is interesting to observe that our methods are essentially topological, except for some very elementary details. The principal contributions of the present paper, as far as new results are concerned, are contained in the main theorem in 2.12 and in the fundamental lemma in 2.1.

0.8. As far as the general plan of approach is concerned, the present study is based upon the work of Geöcze (see bibliography). However, our results are more comprehensive, and our statements and methods are far less involved. We shall make extensive use of two fundamental topological conceptions which were not available to Geöcze: first, the conception of *the topological index*, or order, of a point with respect to a plane closed continuous path-curve, and second, the conception of a *cyclic element* of a Peano space. The application of the latter to our problem was suggested by the work of Morrey who was the first to apply the structure theory of Peano spaces (see Kuratowski-Whyburn [1], also for further references) in the theory of the area of surfaces (see Morrey [1]). While the topological index has been a standard tool in the theory of the area for many years, various useful modifications were suggested to the author by the work of Eilenberg on the topology of the plane (Eilenberg [1]).

0.9. For the convenience of the reader, we collect here some of the most frequently used symbols and terms, to avoid excessive cross-references.

SYMBOLS. U, 0.2; $E(S)$, 0.3; $E(\mathfrak{P})$, 0.4; $A(S)$, 0.5, 0.6; $d(S, \mathfrak{P})$, 0.6; Σ, 1.1; ω, 1.20.

TERMS. Continuous path-surface, 0.2; polyhedral path-surface, 0.3; polyhedron, 0.4; non-degenerate vector-function, 1.3; single-valued continuous argument, 1.5; condition (Arg), 1.6; indicator continuum, 1.7; indicator curve, 1.8; associated Peano space, 1.16; dendrite, 1.19.

In the preceding list, each item is followed by the number of the section where that item is explained.

CHAPTER I. PRELIMINARIES

1.1. In the sequel, only metric spaces will be needed. We shall use Σ to designate a *Peano space*, that is a space which is compact, connected and locally connected.

1.2. Let $x_1(P)$, $x_2(P)$, $x_3(P)$ be three continuous real-valued functions, where the point P varies on a Peano space Σ. We shall designate by $\mathfrak{x}(P)$ the continuous vector-function with components $x_1(P)$, $x_2(P)$, $x_3(P)$, and we shall write

$$\mathfrak{x}(P) = (x_1(P), x_2(P), x_3(P)), \qquad P \in \Sigma.$$

1.3. The continuous vector-function $\mathfrak{x}(P)$ is *non-degenerate* on Σ if it is not constant on any continuum containing more than one point.

1.4. We shall associate with the continuous vector-function $\mathfrak{x}(P)$ three continuous complex-valued functions as follows:

$$f_1(P) = x_2(P) + ix_3(P),$$
$$f_2(P) = x_3(P) + ix_1(P),$$
$$f_3(P) = x_1(P) + ix_2(P).$$

1.5. Let $f(P)$ be a continuous complex-valued function on the Peano-space Σ. Let C be a continuum on Σ, such that $f(P) \neq 0$ on C. If there exists on C a single-valued, continuous, real-valued function $\varphi(P)$ such that

$$f(P) = |f(P)| (\cos \varphi(P) + i \sin \varphi(P)), \qquad P \in C,$$

then $\varphi(P)$ will be called a *single-valued continuous argument* for $f(P)$ on C. The following facts are well known.

1) If $\varphi(P)$, $\psi(P)$ are two single-valued continuous arguments for $f(P)$ on C, then on C the difference $\varphi(P) - \psi(P)$ is equal to a constant of the form $2k\pi$, k an integer.

2) If $\varphi(P)$ is a single-valued continuous argument for $f(P)$ on C, and if k is an integer, then $\varphi(P) + 2k\pi$ is also a single-valued continuous argument for $f(P)$ on C.

3) If C is an arc (topological image of the interval $0 \leqq t \leqq 1$), and if $f(P) \neq 0$ on C, then $f(P)$ possesses a single-valued continuous argument on C.

1.6. Let $f(P)$ be a complex-valued continuous function on a Peano space Σ, and let C be a continuum on Σ. We shall say that $f(P)$ satisfies condition (Arg) on C, if for every choice of the complex constant ζ, such that $f(P) \neq \zeta$ on C, the function $f(P) - \zeta$ possesses a single-valued continuous argument on C. Here and in the sequel, the term *complex* is used to mean *complex or real*, as usual.

1.7. Let $\mathfrak{x}(P)$ be a continuous vector-function on a Peano space Σ. A continuum C on Σ will be called an *indicator continuum* for $\mathfrak{x}(P)$, if at least one of the associated functions $f_1(P)$, $f_2(P)$, $f_3(P)$, defined in 1.4, fails to satisfy condition (Arg) on C.

1.8. If an indicator continuum is a simple closed curve (topological image of the perimeter of the unit circle), then C will be called an *indicator curve* for $\mathfrak{x}(P)$.

1.9. The definition of an indicator curve may be re-stated in the following equivalent form. Let C be a simple closed curve on Σ. Let us assign a positive direction on C. The equation

$$\mathfrak{x} = \mathfrak{x}(P), \qquad P \in C$$

defines then a *closed continuous path-curve* C' in (x_1, x_2, x_3) space. Let us introduce the complex variables

$$\zeta_1 = x_2 + ix_3, \qquad \zeta_2 = x_3 + ix_1, \qquad \zeta_3 = x_1 + ix_2.$$

If we project C' orthogonally upon the ζ_1, ζ_2, ζ_3 planes respectively, we obtain three plane closed continuous path-curves C_1', C_2', C_3', which are represented by the equations (cf. 1.4)

$$C_1': \quad \zeta_1 = f_1(P), \qquad P \in C,$$
$$C_2': \quad \zeta_2 = f_2(P), \qquad P \in C,$$
$$C_3': \quad \zeta_3 = f_3(P), \qquad P \in C.$$

Let ζ_i^0 be a point of the plane ζ_i. Then the *topological index* (or order) of ζ_i^0 with respect to C_i' is defined as follows. If ζ_i^0 is on C_i', then the index is zero. If ζ_i^0 is not on C_i', then the continuously varying argument of the function $f_i(P) - \zeta_i^0$ changes by an amount $2n\pi$, n an integer, while P describes C in the positive sense. The integer n is then the index of ζ_i^0 with respect to C_i' (if we reverse the positive direction on C, then the index changes merely its sign). The index, with respect to C_i', is thus a function defined for all points of the plane ζ_i. Let us designate it by $\mu_i(\zeta_i)$. This function really depends also upon $\mathfrak{x}(P)$ and C, but it is unnecessary for our purposes to use more complicated notations.

Clearly then, the simple closed curve C is an indicator curve for $\mathfrak{x}(P)$, in the sense of 1.8, if and only if at least one of the index-functions $\mu_1(\zeta_1)$, $\mu_2(\zeta_2)$, $\mu_3(\zeta_3)$ is not identically equal to zero.

1.10. LEMMA. *On a Peano space Σ, let there be given a continuous complex-valued function $f(P)$. Suppose there exists a single-valued continuous argument for $f(P)$ on every simple closed curve on which $f(P) \neq 0$. Then there also exists a single-valued continuous argument for $f(P)$ on every domain (connected open set) on which $f(P) \neq 0$.*

PROOF. Let D be any domain on which $f(P) \neq 0$. In D we pick a point P_0, and an argument φ_0 of the complex number $f(P_0)$. We shall keep P_0 and φ_0 fixed. Let P_1 be any point of D different from P_0, and let γ be any arc that joins P_0 and P_1 in D. On γ we have a single-valued continuous argument $\varphi(P;\gamma)$ of $f(P)$, such that $\varphi(P_0;\gamma) = \varphi_0$ (cf. 1.5). Clearly, our lemma is proved if we can show that

$$\varphi(P_1;\gamma') = \varphi(P_1;\gamma'') \tag{5}$$

for any two arcs γ'; γ'' which join P_0 and P_1 in D.

To prove (5), we denote by E the set of all common points of γ' and γ''. Then E is a closed subset of γ' which contains (at least) the points P_0 and P_1. We consider the function

$$\psi(P) = \frac{\varphi(P;\gamma'') - \varphi(P;\gamma')}{2\pi}, \qquad P \in E.$$

Then $\psi(P)$ is well-defined and continuous on E, and it assumes only integral values on E. Indeed, for every point $P \in E$, $\varphi(P;\gamma'')$ and $\varphi(P;\gamma')$ are arguments

of the same complex number, namely, of $f(P)$. Hence, $\psi(P)$ is an integer. We have

$$\psi(P_0) = 0.$$

Let us assume now, in contradiction with (5), that

$$\psi(P_1) \neq 0.$$

Let us designate by E_0, E_1 the subsets of E on which $\psi(P) = 0$, $\psi(P) \neq 0$ respectively. From the properties of $\psi(P)$ described above, it follows that E_0, E_1 are two non-vacuous closed sets on γ' without common points. It follows that we have a sub-arc γ_0' of γ', bounded by two (distinct) points Q_0, Q_1, such that

$$Q_0 \in E_0, \qquad Q_1 \in E_1, \qquad E \cdot \gamma_0' = Q_0 + Q_1. \tag{6}$$

These same two points Q_0, Q_1 are the endpoints of a sub-arc γ_0'' of γ'', and from (6) it follows that

$$\gamma_0' \cdot \gamma_0'' = Q_0 + Q_1. \tag{7}$$

Hence γ_0' and γ_0'' form a simple closed curve C_0 on which $f(P) \neq 0$, since $C_0 \subset D$. By assumption, we have a single-valued continuous argument $\varphi(P)$ for $f(P)$ on C_0. We have then (cf. 1.5)

$$\varphi(P) - \varphi(P; \gamma') = \text{constant on } \gamma_0',$$
$$\varphi(P) - \varphi(P; \gamma'') = \text{constant on } \gamma_0''.$$

Hence

$$\varphi(Q_0) - \varphi(Q_0; \gamma') = \varphi(Q_1) - \varphi(Q_1; \gamma'),$$
$$\varphi(Q_0) - \varphi(Q_0; \gamma'') = \varphi(Q_1) - \varphi(Q_1; \gamma'').$$

Subtraction yields

$$\psi(Q_0) = \psi(Q_1).$$

This is a contradiction, since $Q_0 \in E_0$, $Q_1 \in E_1$.

1.11. LEMMA. *On a Peano space Σ, let there be given a continuous complex-valued function $f(P)$. Suppose there exists a single-valued continuous argument for $f(P)$ on every simple closed curve on which $f(P) \neq 0$. Then there also exists a single-valued continuous argument for $f(P)$ on every continuum on which $f(P) \neq 0$.*

PROOF. Let Γ be any continuum on which $f(P) \neq 0$. Since Σ is locally connected and since $f(P)$ is continuous, there exists, for every point $Q \in \Gamma$, a connected open set $D(Q)$, containing Q, on which $f(P) \neq 0$. Let D denote the sum of all the sets $D(Q)$ which are associated in this manner with the points Q of Γ. Then D is a domain, and $f(P) \neq 0$ on D. By 1.10, we have a single-valued continuous argument for $f(P)$ on D and hence also on Γ, since $\Gamma \subset D$.

1.12. LEMMA. *Let $\mathfrak{x}(P)$ be a continuous vector function on a Peano space Σ. If there exists an indicator continuum for $\mathfrak{x}(P)$ on Σ, then there also exists an indicator curve for $\mathfrak{x}(P)$ on Σ.*

This lemma is a direct consequence of 1.11.

1.13. Let $\mathfrak{x}(P)$ be a continuous vector-function on a Peano space Σ. Then Σ is the sum of continua, without common points, on each of which $\mathfrak{x}(P)$ is constant, and each of which is maximal with respect to this property. Let K designate the class of these continua. Then, in the terminology of R. L. Moore, K is an upper semi-continuous collection of continua, filling up Σ. A comprehensive theory of such collections is contained in Moore [1], to which the reader is referred for proofs of the facts listed below. Each point P of Σ is on precisely one continuum $\Gamma \in K$. Let us write $T(P) = \Gamma$ for the transformation which carries each point P of Σ into that continuum $\Gamma \in K$ which contains P. The class K can be topologized in such a way that this transformation $T(P) = \Gamma$ becomes continuous. K is then again a Peano space. For greater clarity, we shall use Σ^* to denote the space that is obtained in this manner from K, and we shall use P^* to refer to a generic point of Σ^*.

1.14. The transformation T, defined in 1.13, may now be written in the form

$$T(P) = P^*, \qquad P \in \Sigma, \qquad P^* \in \Sigma^*.$$

$T(P)$ is single-valued and continuous on Σ. The inverse transformation $T^{-1}(P^*)$ need not be single-valued on Σ^*, but for every point P^* of Σ^* the inverse set $T^{-1}(P^*)$ is a continuum, namely one of the continua $\Gamma \in K$. The following statements are consequences of the preceding statements (see Moore [1]).

a) If C^* is a continuum on Σ^*, then $T^{-1}(C^*)$ is a continuum on Σ.

b) If D^* is a domain (connected open set) on Σ^*, then $T^{-1}(D^*)$ is a domain on Σ.

c) If F^* is a closed set on Σ^*, then the set $\Sigma^* - F^*$ has the same number of components as the set $\Sigma - T^{-1}(F^*)$. In fact, the components of $\Sigma - T^{-1}(F^*)$ are simply the inverse sets of the components of $\Sigma^* - F^*$.

1.15. By means of the transformation $T(P) = P^*$, the continuous vector-function $\mathfrak{x}(P)$ is carried into a continuous vector-function $\mathfrak{x}^*(P^*)$ in the following sense. Given $P^* \in \Sigma^*$, let us put

$$\mathfrak{x}^*(P^*) = \mathfrak{x}(P), \qquad P \in T^{-1}(P^*). \tag{8}$$

Since $\mathfrak{x}(P)$ is constant on $T^{-1}(P^*)$, the relation (8) defines a single-valued and obviously continuous vector function on Σ^*. In a similar manner, the components $x_i(P)$ of $\mathfrak{x}(P)$, as well as the associated functions $f_i(P)$ defined in 1.4, are transformed by T into single-valued continuous functions on Σ^* which we shall denote by $x_i^*(P^*)$, $f_i^*(P^*)$ respectively. In a general way, if $g(P)$ is any continuous function on Σ, such that $g(P)$ is constant on each continuum $\Gamma \in K$, then the formula

$$g^*(P^*) = g(P), \qquad P \in T^{-1}(P^*)$$

transforms $g(P)$ into a single-valued continuous function $g^*(P^*)$ on Σ^*.

1.16. We shall call Σ^* *the associated Peano space* and the functions $\mathfrak{x}^*(P^*)$, $x_i^*(P^*)$, $f_i^*(P^*)$, $g^*(P^*)$ *the associated functions*, with respect to $\mathfrak{x}(P)$.

1.17. LEMMA. *Let $\mathfrak{x}(P)$ be a continuous vector function on a Peano space Σ. Let $f^*(P^*)$ be a continuous complex-valued function on the associated Peano space Σ^*, and let C^* be a continuum on Σ^*, such that $f^*(P^*)$ fails to satisfy condition* (Arg) *on C^*. Then the function*

$$f(P) = f^*(T(P))$$

fails to satisfy condition (Arg) *on the continuum $C = T^{-1}(C^*)$.*

PROOF. Obviously $f(P)$ is a single-valued continuous function on Σ, while $C = T^{-1}(C^*)$ is a continuum by 1.14. Assume now, in contradiction to the lemma, that $f(P)$ satisfies condition (Arg) on C. Let then ζ be any complex constant such that

$$f^*(P^*) - \zeta \neq 0 \quad \text{on} \quad C^*. \tag{9}$$

We have then also

$$f(P) - \zeta \neq 0 \quad \text{on} \quad C.$$

By assumption, we have therefore a single-valued continuous argument $\varphi(P)$ for $f(P) - \zeta$ on C. We assert that the formula

$$\varphi^*(P^*) = \varphi(P), \qquad P \in T^{-1}(P^*), \qquad P^* \in C^*,$$

defines a single-valued continuous argument for $f^*(P^*)$ on C^*. Clearly, we only have to show that

$$\varphi(P') = \varphi(P'')$$

for any two points P', P'' such that

$$P' \in T^{-1}(P^*), \qquad P'' \in T^{-1}(P^*), \qquad P^* \in C^*.$$

Now $T^{-1}(P^*)$ is a continuum Γ comprised in C, such that $f(P)$ is constant on Γ. Hence, $\varphi(P)$ is defined and continuous on $T^{-1}(P^*)$. Furthermore, for each $P \in T^{-1}(P^*)$, $\varphi(P)$ is an argument of the same complex number $f(P) = f^*(T(P)) = f^*(P^*)$. Hence the set of values of $\varphi(P)$ on $T^{-1}(P^*)$ is denumerable. As $T^{-1}(P^*)$ is a continuum and $\varphi(P)$ is continuous on $T^{-1}(P^*)$, it follows that $\varphi(P)$ is constant on $T^{-1}(P^*)$. Hence $\varphi(P') = \varphi(P'')$ for P', $P'' \in T^{-1}(P^*)$, $P^* \in C^*$. Since ζ was arbitrary, except for the condition (9), we reached a contradiction to the assumption that $f^*(P^*)$ fails to satisfy condition (Arg) on C^*, and the lemma is proved.

1.18. LEMMA. *Let $\mathfrak{x}(P)$ be a continuous vector-function on a Peano space Σ. Let $\mathfrak{x}^*(P^*)$ be the associated vector-function on the associated Peano space Σ^*. If there exists an indicator continuum for $\mathfrak{x}^*(P^*)$ on Σ^*, then there also exists an indicator continuum for $\mathfrak{x}(P)$ on Σ.*

PROOF. If C^* is an indicator continuum for $\mathfrak{x}^*(P^*)$ on Σ^*, then $C = T^{-1}(C^*)$ is an indicator continuum for $\mathfrak{x}(P)$ on Σ, as a direct consequence of 1.17.

1.19. A Peano space is a *dendrite* if it possesses no proper cyclic element. The following lemma is an easy consequence of the general theorem concerning

the structure of a Peano space in terms of cyclic chains (see Kuratowski-Whyburn [1]). Hence the details of the proof will be left to the reader.

LEMMA. *Suppose the Peano space Σ is a dendrite. Then for every $\delta > 0$ there exists on Σ a finite system of points $Q_1, Q_2, \cdots, Q_n$ with the following properties.*

1) *For $k = 1, 2, \cdots, n$, the set $\Sigma - Q_k$ has precisely two components.*

2) *Each component of the set $\Sigma - (Q_1 + \cdots + Q_n)$ has a diameter less than δ.*

1.20. If $g(P)$ is a real or complex-valued function on a Peano space Σ, and E is any subset of Σ, then $\omega(g(P); E)$ will denote the oscillation of $g(P)$ on E. That is

$$\omega(g(P); E) = \text{l.u.b.} \mid g(P'') - g(P') \mid, \qquad P', P'' \in E.$$

Similarly, if $\mathfrak{x}(P)$ is a continuous vector-function on Σ, its oscillation on E is given by

$$\omega(\mathfrak{x}(P); E) = \text{l.u.b.} \mid \mathfrak{x}(P'') - \mathfrak{x}(P') \mid, \qquad P', P'' \in E,$$

where $\mid \mathfrak{x}(P'') - \mathfrak{x}(P') \mid$ designates the length of the vector $\mathfrak{x}(P'') - \mathfrak{x}(P')$.

1.21. The following elementary lemma, due to Geöcze, is very important both in his work and in ours.

LEMMA. *Given a continuous path-surface as in* (3), *suppose that for every $\epsilon > 0$ there exists on U a finite system of simple closed curves $C_1, C_2, \cdots, C_m$ with the following properties.* (*i*) *No two of these curves intersect each other.* (*ii*) *On each of the regions into which U is divided by this system of curves, the oscillation of $\mathfrak{x}(P)$ is less than ϵ. Then $A(S) = 0$.*

PROOF. From (i) it follows that the curves C_j determine on U exactly $m + 1$ regions which we denote, in any order, by $R_1, R_2, \cdots, R_{m+1}$. From (ii) it follows that the oscillation of $\mathfrak{x}(P)$ on each of the curves C_j is also $\leqq \epsilon$. Let us now select on each one of the curves C_j a point P_j, and in each one of the regions R_k an interior point Q_k. We have then

$$\mid \mathfrak{x}(P) - \mathfrak{x}(P_j) \mid \leqq \epsilon \quad \text{for} \quad P \in C_j, \qquad j = 1, 2, \cdots, m, \tag{10}$$

$$\mid \mathfrak{x}(Q) - \mathfrak{x}(Q_k) \mid \leqq \epsilon \quad \text{for} \quad Q \in R_k, \qquad k = 1, 2, \cdots, m + 1. \tag{11}$$

Let us now take a triangulation T of U, such that a) the maximum diameter of the triangles of T is less than ϵ, b) each one of the curves C_j is a sum of sides of triangles of T, and c) no triangle of T has vertices on two different curves C_j. Such a triangulation T obviously exists. We now proceed to define a corresponding polyhedron $\mathfrak{P}$ (see 0.4) as follows. Let V be a vertex of T. We assign the point V' as follows. If V is an interior point of a region R_k, then $V' = \mathfrak{x}(Q_k)$. If V is on a curve C_j, then $V' = \mathfrak{x}(P_j)$. In view of (10), (11) and condition a), it follows that the resulting polyhedron $\mathfrak{P}$ satisfies the condition

$$d(S, \mathfrak{P}) \leqq 2\epsilon. \tag{12}$$

Further, since each triangle of T has either two vertices on the same curve C_j,

or else has at least two vertices interior to the same region R_k, it follows that

$$E(\mathfrak{P}) = 0. \tag{13}$$

Thus, for every $\epsilon > 0$, we have a polyhedron $\mathfrak{P}$ satisfying (12) and (13). By 0.6 it follows that $A(S) = 0$.

1.22. LEMMA. *Given a continuous path-surface S as in* (3), *suppose that there exists an indicator curve for $\mathfrak{x}(P)$ on U. Then $A(S) > 0$.*

This lemma is well known, except for the terminology used here. See for example Radó [1]; the proof given there for Jordan surfaces applies, after obvious modifications, to a general continuous path-surface. In verifying the lemma, the term indicator curve should be used in the equivalent sense described in 1.9.

CHAPTER II. THE MAIN RESULTS

2.1. The fundamental lemma. *Let Σ be a Peano space, such that for every totally disconnected closed set Δ on Σ, the set $\Sigma - \Delta$ is connected. Let the continuous vector-function $\mathfrak{x}(P)$ be non-degenerate on Σ. Then there exists an indicator curve for $\mathfrak{x}(P)$ on Σ.*

PROOF. In view of 1.12, it is sufficient to prove that there exists an indicator continuum for $\mathfrak{x}(P)$ on Σ. Since $\mathfrak{x}(P)$ is non-degenerate, one of its components, say $x_1(P)$, is not constant on Σ. Let m_1, M_1 be the minimum and the maximum of $x_1(P)$ on Σ respectively. We have then $m_1 < M_1$, and there exist two points O, O^* on Σ such that

$$x_1(O) = m_1, \qquad x_1(O^*) = M_1.$$

Put

$$\xi = \frac{m_1 + M_1}{2},$$

and consider on Σ the set

$$E = \underset{P}{E}[x_1(P) < \xi].$$

That is, E consists of those points P of Σ for which $x_1(P) < \xi$. Then E is a non-vacuous open set, and we have

$$O \in E, \qquad O^* \in \Sigma - \bar{E}$$

where $\bar{E}$ designates the closure of E. Let D be the component of E which contains O. Then D is a domain containing O. The frontier $\mathcal{F}(D)$ of D is a closed set which separates O and O^*. Hence, by our assumption concerning the space Σ, $\mathcal{F}(D)$ is not totally disconnected. Hence $\mathcal{F}(D)$ has some component C_0 which is a continuum not reducing to a single point. Since $\mathfrak{x}(P)$ is non-degenerate by assumption, and since $x_1(P) \equiv \xi$ on C_0, at least one of $x_2(P)$, $x_3(P)$, say $x_2(P)$, is not constant on C_0. Since C_0 is compact, we have therefore on C_0 two points Q_0, Q_0^* such that

$$x_2(Q_0) = m_2, \qquad x_2(Q_0^*) = M_2,$$

where m_2, M_2 designate the minimum and the maximum respectively of $x_2(P)$ on C_0. Note that

$$m_2 < M_2,$$

since $x_2(P)$ is not constant on C_0.

Since Σ is locally connected and $\mathfrak{x}(P)$ is continuous, we have two continua κ_0, κ_0^*, containing Q_0, Q_0^* as interior points respectively, and so small in diameter that

$$\omega(x_1(P), \kappa_0), \qquad \omega(x_1(P), \kappa_0^*) < \frac{M_2 - m_2}{4},$$

$$\omega(x_2(P), \kappa_0), \qquad \omega(x_2(P), \kappa_0^*) < \frac{M_2 - m_2}{4}.$$

As Q_0, Q_0^* are points of the frontier $\mathfrak{F}(D)$ of D, there exist points A_0, A_0^* such that

$$A_0 \in \kappa_0 \cdot D, \qquad A_0^* \in \kappa_0^* \cdot D.$$

Let γ be an arc that joins A_0 and A_0^* in D. As γ is a closed subset of D and $x_1(P) < \xi$ on D, we have a number ξ_0 such that

$$x_1(P) \leqq \xi_0 < \xi \quad \text{for} \quad P \in \gamma.$$

Let us put

$$\zeta_3 = \frac{\xi_0 + \xi}{2} + i\,\frac{M_2 + m_2}{2}.$$

2.2. We shall use the function (cf. 1.4)

$$f_3(P) = x_1(P) + ix_2(P)$$

to show presently that the set

$$\Gamma = C_0 + \kappa_0 + \kappa_0^* + \gamma$$

is an indicator continuum for $\mathfrak{x}(P)$ on Σ. This will be achieved by establishing the following facts.

a) Γ is a continuum.

b) $f_3(P) - \zeta_3 \neq 0$ on Γ.

c) $f_3(P) - \zeta_3$ does not possess a single-valued continuous argument on Γ.

2.3. Now (a) in 2.2 is obvious, since each one of the sets C_0, κ_0, κ_0^*, γ is a continuum, and

$$C_0\kappa_0 \neq 0, \qquad \kappa_0\gamma \neq 0, \qquad \gamma\kappa_0^* \neq 0, \qquad \kappa_0^* C_0 \neq 0.$$

2.4. Next we show that

$$\Re(f_3(P) - \zeta_3) > 0 \quad \text{on} \quad C_0, \tag{14}$$

$$\Im(f_3(P) - \zeta_3) < 0 \quad \text{on} \quad \kappa_0, \tag{15}$$

$$\Im(f_3(P) - \zeta_3) > 0 \quad \text{on} \quad \kappa_0^*, \tag{16}$$

$$\Re(f_3(P) - \zeta_3) < 0 \quad \text{on} \quad \gamma. \tag{17}$$

In these inequalities, $\Re$ and $\Im$ denote real part and imaginary part respectively, as usual. To verify these relations, observe that we have on C_0

$$\Re(f_3(p) - \zeta_3) = \xi - \frac{\xi_0 + \xi}{2} = \frac{\xi - \xi_0}{2} > 0.$$

On κ_0 we have

$$\begin{aligned}\Im(f_3(P) - \zeta_3) &= x_2(P) - \frac{M_2 + m_2}{2} \\ &\leqq x_2(Q_0) + \omega(x_2(P); \kappa_0) - \frac{M_2 + m_2}{2} \\ &\leqq m_2 + \frac{M_2 - m_2}{4} - \frac{M_2 + m_2}{2} = \frac{m_2 - M_2}{4} < 0.\end{aligned}$$

On κ_0^* we have

$$\begin{aligned}\Im(f_3(P) - \zeta_3) &= x_2(P) - \frac{M_2 + m_2}{2} \\ &\geqq x_2(Q_0^*) - \omega(x_2(P); \kappa_0^*) - \frac{M_2 + m_2}{2} \\ &\geqq M_2 - \frac{M_2 - m_2}{4} - \frac{M_2 + m_2}{2} = \frac{M_2 - m_2}{4} > 0.\end{aligned}$$

On γ we have

$$\begin{aligned}\Re(f_3(P) - \zeta_3) &= x_1(P) - \frac{\xi_0 + \xi}{2} \\ &\leqq \xi_0 - \frac{\xi_0 + \xi}{2} = \frac{\xi_0 - \xi}{2} < 0.\end{aligned}$$

2.5. Condition (b) in 2.2 is an immediate consequence of (14), (15), (16), (17) in 2.4.

2.6. Let us note the following inequalities.

$$\Re(f_3(Q_0) - \zeta_3) > 0, \tag{18}$$

$$\Im(f_3(Q_0) - \zeta_3) < 0, \tag{19}$$

$$\Re(f_3(Q_0^*) - \zeta_3) > 0, \tag{20}$$

$$\Im(f_3(Q_0^*) - \zeta_3) > 0, \tag{21}$$

(18) and (20) follow directly from (14), while (19), (21) follow from (15) and (16) respectively.

2.7. Assume now that $f_3(P) - \zeta_0$ possesses a single-valued continuous argument $\varphi(P)$ on Γ, in contradiction to our assertion (c) in 2.2. By (18), (19) we can assume then that

$$-\frac{\pi}{2} < \varphi(Q_0) < 0. \tag{22}$$

Consider now $\varphi(P)$ on the continuum C_0. We assert that

$$-\pi < \varphi(P) < \pi \quad \text{on} \quad C_0.$$

Otherwise, by virtue of its continuity and in view of (22), $\varphi(P)$ would have to assume at some point of C_0 one of the values $-\pi$, π, which is however impossible by (14). In particular it follows that

$$-\pi < \varphi(Q_0^*) < \pi. \tag{23}$$

In view of (20), (21) it follows from (23) that

$$0 < \varphi(Q_0^*) < \frac{\pi}{2}. \tag{24}$$

Consider now $\varphi(P)$ on the continuum $\kappa_0 + \gamma + \overset{*}{\kappa_0}$. We assert that

$$-2\pi < \varphi(P) < 0 \quad \text{on} \quad \kappa_0 + \gamma + \overset{*}{\kappa_0}.$$

Otherwise, by virtue of its continuity and in view of (22), $\varphi(P)$ would have to assume on $\kappa_0 + \gamma + \overset{*}{\kappa_0}$ one of the values 0, -2π, which is however impossible by (15), (16), (17). In particular, it follows that

$$-2\pi < \varphi(Q_0^*) < 0. \tag{25}$$

Since (24) and (25) obviously contradict each other, the assertion (c) in 2.2 is established, and the proof of the fundamental lemma in 2.1 is complete.

2.8. THEOREM (cf. 0.7). *Let S be a continuous path-surface given as in* (3). *If $\mathfrak{x}(P)$ is non-degenerate on U, then $A(S) > 0$.*

PROOF. It is well known that U is not disconnected by any one of its totally disconnected closed subsets. Hence, by 2.1, there exists an indicator curve for $\mathfrak{x}(P)$ on U. By 1.22 it follows that $A(S) > 0$.

2.9. THEOREM (of Geöcze). *If the representation* (3) *defines a bi-unique correspondence between the points P of U and the points (x_1, x_2, x_3) of S, then $A(S) > 0$.* Briefly: *the Lebesgue area of a simple continuous path-surface is always positive.*

This theorem is of course a special case of 2.8.

2.10. LEMMA. *Let S be a continuous path-surface given as in* (3). *Suppose there exists no indicator curve for $\mathfrak{x}(P)$ on U. Then the associated Peano space Σ^* reduces to a dendrite.*

PROOF. By a well-known theorem of R. L. Moore [2], Σ^* is a *cactoid*. That is, either Σ^* is a dendrite, or else every proper cyclic element of Σ^* is homeomorphic to the surface of the unit sphere. Consider now the associated vector-function $\mathfrak{x}^*(P^*)$ defined in 1.16. By 1.12, there exists no indicator continuum

for $\mathfrak{x}(P)$ on U. By 1.18, it follows that there exists no indicator continuum for $\mathfrak{x}^*(P^*)$ on Σ^*. Since $\mathfrak{x}^*(P^*)$ is non-degenerate on Σ^*, and since by the theorem of R. L. Moore, referred to above, no proper cyclic element of Σ^* is disconnected by any one of its closed totally disconnected subsets, it follows from the fundamental lemma in 2.1, applied to an hypothetical proper cyclic element of Σ^*, that Σ^* cannot have proper cyclic elements, and therefore reduces to a dendrite, as asserted.

2.11. LEMMA. Given a continuous path-surface S as in (3), suppose that the associated Peano space Σ^* reduces to a dendrite. Then for every $\epsilon > 0$ there exists on U a finite system of simple closed curves $C_1, C_2, \cdots, C_m$ with the following properties. (i) No two of the curves $C_1, C_2, \cdots, C_m$ intersect each other. (ii) On each component of the set $U - (C_1 + C_2 + \cdots + C_m)$ the oscillation of $\mathfrak{x}(P)$ is less than ϵ.

PROOF. Consider again the associated vector-function $\mathfrak{x}^*(P^*)$ on Σ^*. By 1.19, for every $\delta > 0$ we have on Σ^* a finite system of points $P_1^*, P_2^*, \cdots, P_n^*$ with the following properties. (i) For each k, the set $\Sigma^* - P_k^*$ has precisely two components. (ii) Each component of the set $\Sigma^* - (P_1^* + P_2^* + \cdots + P_n^*)$ has a diameter less than δ. Since $\mathfrak{x}^*(P^*)$ is uniformly continuous on Σ^*, we can choose δ so small that the oscillation of $\mathfrak{x}^*(P^*)$ on each component of the set $\Sigma^* - (P_1^* + P_2^* + \cdots + P_n^*)$ is less than ϵ. Let $\Gamma_1, \Gamma_2, \cdots, \Gamma_n$ be the continua which correspond to $P_1^*, P_2^*, \cdots, P_n^*$ on U. Then (cf. 1.14) for each k the set $U - \Gamma_k$ has precisely two components, and on each component of the set $U - (\Gamma_1 + \cdots + \Gamma_n)$ the oscillation of $\mathfrak{x}(P)$ is less than ϵ. By well-known theorems on the topology of the sphere (see for example Newman [1]), the proof can now be completed as follows. The set $U - (\Gamma_1 + \cdots + \Gamma_n)$ has precisely $n + 1$ components, each of which is a domain (connected open set). Let $D_1, D_2, \cdots, D_{n+1}$ be these domains, in any order. Consider one of these domains D_k. There exists in D_k a sequence of Jordan regions R_k^j, $j = 1, 2, \cdots$, each bounded by a finite number of simple closed curves, such that (i) $R_k^j \subset R_k^{j+1}$, and (ii) for every point $P \in D_k$ there exists a j_0 such that $P \in R_k^j$ for $j > j_0$. For fixed large j, U is then the sum of $2n + 1$ Jordan regions, without common interior points, namely of the $(n + 1)$ regions $R_1^j, \cdots, R_n^j, R_{n+1}^j$ and of n further doubly connected Jordan regions $r_1^j, \cdots, r_n^j$, each of which contains precisely one of the continua $\Gamma_1, \cdots, \Gamma_n$ in its interior. These latter regions $r_1^j, \cdots, r_n^j$ are doubly connected as a consequence of the fact that for each k the set $U - \Gamma_k$ has precisely two components. Each r_k^j is bounded by two simple closed curves which are parts of the boundaries of two different regions R_k^j. The oscillation of $\mathfrak{x}(P)$ on R_k^j is less than ϵ, since $R_k^j \subset D_k$. On the other hand, if E_k is any open set containing Γ_k, we shall have $r_k^j \subset E_k$ for j sufficiently large. Since $\mathfrak{x}(P)$ is constant on Γ_k, the open set E_k can be so chosen that the oscillation of $\mathfrak{x}(P)$ on E_k is less than ϵ. Hence, if j is sufficiently large, the system consisting of the boundary curves of all the regions $R_1^j, R_2^j, \cdots, R_{n+1}^j$ will possess the properties required in the lemma.

2.12. THE MAIN THEOREM. *Given a continuous path-surface S as in* (3), *the*

Lebesgue area $A(S)$ of S is equal to zero if and only if there exists no indicator curve for $\mathfrak{x}(P)$ on U.

PROOF. The condition is necessary by 1.22. To prove the sufficiency, assume that there exists no indicator curve for $\mathfrak{x}(P)$ on U. By 2.10, the associated Peano space reduces then to a dendrite. By 2.11, it follows that the assumptions of the lemma in 1.21 are satisfied, and hence $A(S) = 0$.

2.13. THEOREM (cf. 0.7). *The Lebesgue area $A(S)$ of a continuous path-surface S, given as in* (3), *is equal to zero if and only if the associated Peano space Σ^* reduces to a dendrite.*

PROOF. The necessity follows directly from 1.22 and 2.10, while the sufficiency follows from 2.11 and 1.21.

2.14. THEOREM (of Geöcze). *The Lebesgue area $A(S)$ of a continuous path-surface S, given as in* (3), *is equal to zero if and only if for every $\epsilon > 0$ there exists on U a finite system of simple closed curves $C_1, C_2, \cdots, C_m$ with the following properties. (i) No two of these curves intersect each other. (ii) The oscillation of $\mathfrak{x}(P)$ is less than ϵ on each component of the set $U - (C_1 + C_2 + \cdots + C_m)$.*

PROOF. The necessity follows from 2.13 and 2.11, while the sufficiency follows from 1.21.

CHAPTER III. MISCELLANEOUS REMARKS

3.1. Given a continuous path-surface S as in (3), let us denote by $[S]$ the set of points (x_1, x_2, x_3) which correspond to the points P of U by means of (3). We already observed that the area $A(S)$ of S is not determined by the point-set $[S]$, but rather by the manner in which $[S]$ is described by virtue of (3). This remark will be illustrated presently by examples, due essentially to Geöcze.

3.2. Let us denote by M_0 the half-circle on U which passes through the points $(0, 0, -1)$, $(1, 0, 0)$, $(0, 0, 1)$. Let $\mathfrak{x}_0(P)$ be any continuous vector-function defined on M_0. If P is a point of U not on M_0, then let us denote by $\pi(P)$ the small circle on U which passes through P and is contained in a plane parallel to the (x_1, x_2) plane. Let the vector-function $\mathfrak{x}(P)$ be defined on U as follows.

$$\mathfrak{x}(P) = \mathfrak{x}_0(P) \text{ on } M_0,$$

$$\mathfrak{x}(P) = \text{constant on } \pi(P).$$

Then $\mathfrak{x}(P)$ is clearly continuous on U. Consider now the continuous path-surface S determined by this particular vector-function $\mathfrak{x}(P)$. We assert that $A(S) = 0$. Indeed, let T be any triangulation of U, such that each triangle of T has a side which is a sub-arc of a great circle passing through the points $(0, 0, 1)$ and $(0, 0, -1)$. For each vertex V of T, let us assign the point $\mathfrak{x}(V)$ as the point V' of 0.4. We obtain in this manner a polyhedron $\mathfrak{P}$, such that $d(S, \mathfrak{P})$ is equal to the maximum diameter of the triangles of T, and such that $E(\mathfrak{P}) = 0$. By 0.6, it follows that $A(S) = 0$. This was to be expected, since the path-surface S under consideration may be thought of as reducing, in a

way, to the continuous path-curve Γ given by

$$\Gamma: \quad \mathfrak{x} = \mathfrak{x}_0(P), \qquad P \in M_0 . \tag{26}$$

Let us designate by $[\Gamma]$ the set of those points (x_1, x_2, x_3) which correspond to the points P of M_0 by means of (26). Clearly then $[\Gamma] = [S]$. On the other hand, it is well known that by proper choice of $\mathfrak{x}_0(P)$ the point-set $[\Gamma]$ and hence also the point-set $[S]$ can be made to coincide with any desired compact, connected and locally connected point-set in (x_1, x_2, x_3) space.

3.3. In particular, we can make $[S]$ coincide with the cube $-1 \leqq x_i \leqq 1$, $i = 1, 2, 3$. We obtain then the continuous path-surface S of Geöcze [3] which fills a cube and has zero area.

3.4. Next, let us choose $[S]$ as a point-set U' homeomorphic to U, the surface of the unit sphere. The construction described in 3.2 yields then a continuous path-surface S such that $[S] = U'$ and $A(S) = 0$. On the other hand, we have by assumption a continuous vector-function $\bar{\mathfrak{x}}(P)$ on U such that $\bar{\mathfrak{x}}(P_1) \neq \bar{\mathfrak{x}}(P_2)$ for $P_1 \neq P_2$, and such that for the continuous path-surface

$$\bar{S}: \quad \mathfrak{x} = \bar{\mathfrak{x}}(P), \qquad P \in U$$

we have $[\bar{S}] = U'$. By 2.9 we have then $A(\bar{S}) > 0$. Thus we have two continuous path-surfaces

$$S: \quad \mathfrak{x} = \mathfrak{x}(P), \qquad P \in U, \tag{27}$$

$$\bar{S}: \quad \mathfrak{x} = \bar{\mathfrak{x}}(P), \qquad P \in U, \tag{28}$$

with the following properties. (i) $[S] = [\bar{S}] = U'$, where U' is homeomorphic to U. (ii) For each point (x_1, x_2, x_3) on U', we have *precisely one point* $P \in U$ which is carried into (x_1, x_2, x_3) by (28) and we have *at least one point* $P \in U$ which is carried into (x_1, x_2, x_3) by (27). (iii) Finally, contrarily to what one may have expected under such circumstances, we have $A(S) < A(\bar{S})$, since $A(S) = 0$, $A(\bar{S}) > 0$.

3.5. In a more picturesque way, we can describe this situation as follows. Let U' be a topological image of the surface U of the unit sphere. Then the area of a continuous path-surface S, given as in (3), is certainly positive if S covers U' *just once*. On the other hand, if S is required to cover all of U', but is permitted to cover *parts of* U' *several times*, then the area of S may be equal to zero. Of course, there is nothing really paradoxical about this, but such a situation is an illustration of the fundamental differences between the theory of the length of curves and the theory of the area of surfaces.

3.6. We pass on to remarks concerning certain interesting aspects of our main theorem in 2.12. A *closed continuous path-curve* Γ, in (x_1, x_2, x_3) space, is determined by an equation

$$\Gamma: \quad \mathfrak{x} = \mathfrak{x}(P), \qquad P \in C, \tag{29}$$

where C is a simple closed curve and $\mathfrak{x}(P)$ is a continuous vector-function on C. It is assumed that a positive direction is assigned on C. Let p be any plane.

Let P' be the point (x_1, x_2, x_3) which corresponds to the point P of C by means of (29), and let P_p be the orthogonal projection of P' upon the plane p. Let finally $\mathfrak{x}_p(P)$ be the vector that joins the point (0, 0, 0) to the point P_p. Then we shall designate by Γ_p the closed continuous path-curve

$$\Gamma_p: \qquad \mathfrak{x} = \mathfrak{x}_p(P), \qquad P \in C.$$

We shall call Γ_p *the orthogonal projection* of Γ upon the plane p. Let Q_p be a variable point in the plane p. We define then the index-function $\mu(Q_p, \Gamma_p)$ as follows. If Q_p is on Γ_p, then $\mu(Q_p, \Gamma_p) = 0$. Otherwise, $\mu(Q_p, \Gamma_p)$ is equal to the topological index of Q_p with respect to Γ_p. For conciseness, we shall use the subscripts 1, 2, 3 to refer to the coordinate planes (x_2, x_3), (x_3, x_1), (x_1, x_2) respectively. Then the index-functions $\mu(Q_i, \Gamma_i)$, $i = 1, 2, 3$, are identical with the index-functions $\mu_i(\zeta_i)$ of 1.9.

3.7. Consider now a continuous path-surface

$$S: \qquad \mathfrak{x} = \mathfrak{x}(P), \qquad P \in U. \tag{30}$$

If C is any simple closed curve on U, then there corresponds to C, by means of (30), a closed continuous path-curve

$$\Gamma: \qquad \mathfrak{x} = \mathfrak{x}(P), \qquad P \in C. \tag{31}$$

Suppose now that there exists no indicator curve for $\mathfrak{x}(P)$ on U. By our main theorem in 2.12, $A(S) = 0$. Since $A(S)$ is invariable under changes of the Cartesian coordinate system, it follows (cf. 1.22, 1.9, 3.6) that $\mu(Q_p, \Gamma_p) \equiv 0$ for every choice of the plane p. Hence we have the

THEOREM. *If for every closed continuous path-curve* Γ *on a continuous path-surface* S, *given as in* (30) *and* (31) *respectively, the three index-functions* $\mu_i(\zeta_i)$, $i = 1, 2, 3$, *are identically equal to zero, then the index-function* $\mu(Q_p, \Gamma_p)$ *is also identically equal to zero for every choice of the plane* p.

3.8. One may be tempted to strengthen the preceding statement as follows: If for a closed continuous path-curve Γ, given as in (29), the three index-functions $\mu(Q_i, \Gamma_i)$, $i = 1, 2, 3$, are identically equal to zero, then the index-function $\mu(Q_p, \Gamma_p)$ is also identically equal to zero, for every choice of the plane p. *This statement is however false.* A first counter-example has been suggested to the writer by J. v. Neumann. Subsequently, W. Scott, a student of the writer, found an extremely simple construction for examples serving various further purposes also.

3.9. In conclusion, we shall give a brief description of the condition used by Geöcze in his characterization of continuous path-surfaces of zero area. Geöcze assumes that the vector-function $\mathfrak{x}(P)$ which determines the surface is defined on the unit square. Hence, his continuous path-surface is given by an equation

$$S: \qquad \mathfrak{x} = \mathfrak{x}(u, v), \qquad (u, v) \in s_0, \tag{32}$$

where

$$s_0: \qquad 0 \leqq u \leqq 1, \qquad 0 \leqq v \leqq 1. \tag{33}$$

In stating the definitions of Geöcze, we shall use present-day terminology if possible; otherwise, we shall use literal translations of the terms invented by Geöcze. All point-sets occurring in 3.10 to 3.15 are in the (u, v) plane unless stated otherwise.

3.10. A *chain* consists of a finite number of straight segments placed end to end. These segments are permitted to intersect each other.

3.11. We recall that a domain D is a connected open set. The frontier $\mathcal{F}(D)$ of D satisfies the relation $\mathcal{F}(D) = \bar{D} - D$, where $\bar{D}$ is the closure of D.

3.12. If D is a simply connected domain, then I will denote the set of those points of $\mathcal{F}(D)$ which are accessible from within D by some chain. The points of I can be cyclicly ordered.

3.13. A domain D is *simple* if the following conditions hold. (i) D is bounded. (ii) Let D_∞ denote the unbounded component of the complement of $\bar{D}$. Then $\mathcal{F}(D) = \mathcal{F}(D_\infty)$. Geöcze notes that a simple domain is also simply connected.

3.14. A *four-sided domain* is a figure consisting of a *simple* domain (see 3.13) and of four point-sets α_1, α_2, α_3, α_4, called the *sides*, satisfying the following conditions. (i) α_1, α_2, α_2, α_4 are subsets of the set I defined in 3.12. (ii) $\bar{\alpha}_1\bar{\alpha}_3 = 0$, $\bar{\alpha}_2\bar{\alpha}_4 = 0$, where the bars denote closure. (iii) $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = I$. (iv) If P_1, P_2, P_3, P_4 are any four points such that $P_i \in \alpha_i$, $i = 1, 2, 3, 4$, then the cyclic order of these points, on the set I, is P_1, P_2, P_3, P_4.

3.15. Given the continuous path-surface (32), the symbol $(\Phi, \Psi, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$ will be meaningful if the following conditions are satisfied. (i) Φ and Ψ are two different ones of the components $x_1(u, v)$, $x_2(u, v)$, $x_3(u, v)$, of the vector-function $x(u, v)$ appearing in (32). (ii) λ_1, λ_2, λ_3, λ_4 are real numbers such that $\lambda_3 > \lambda_1$, $\lambda_4 > \lambda_2$. (iii) There exists in the unit square s_0 a *four-sided domain*, with sides α_1, α_2, α_3, α_4, such that $\Phi = \lambda_1$ on α_1, $\Phi = \lambda_3$ on α_3, $\Psi = \lambda_2$ on α_2, $\Psi = \lambda_4$ on α_4. Such a four-sided domain will then be denoted by the symbol $(\Phi, \Psi, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

3.16. THEOREM OF GEÖCZE (see Geöcze [4]). *The Lebesgue area $A(S)$ of the continuous path-surface S, given by* (32), *is equal to zero if and only if there exists no four-sided domain* $(\Phi, \Psi, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

3.17. It is easy to see that the existence of a domain $(\Phi, \Psi, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$ implies the existence of an indicator curve in our sense (see 1.8, 1.9; in the definitions given there, U should of course be replaced now by s_0). But the present writer was unable to derive a proof of reasonable directness and simplicity for the converse statement and hence for the equivalence of the Geöcze condition (see 3.15) with our condition (see 2.12).

THE OHIO STATE UNIVERSITY

BIBLIOGRAPHY

EILENBERG, S. [1] *Transformations continues en circonférences et la topologie du plan*, Fundamenta Mathematicae, vol. 26, 1936, pp. 61–112.

GEÖCZE, Z. [1] *On the theory of the area* (in Hungarian), Mathematikai és természettudományi értesitő, vol. 31, 1913, pp. 306–318.

——— [2] *On the surface of zero area* (in Hungarian), ibid. vol. 33,1915, pp. 730–748.

GEÖCZE, Z. [3] *Example of a surface of zero area which fills a cube* (in Hungarian), Mathematikai és physikai lapok, vol. 23, 1914, pp. 115–117.

Recherches générales sur la quadrature des surfaces courbes, Part 1, Mathematische und naturwissenschaftliche Berichte aus Ungarn, vol. 27, 1909, pp. 1–21.

KURATOWSKI, C., AND WHYBURN, G. T. [1] *Sur les éléments cyclic et leurs applications*, Fundamenta Mathematicae, vol. 16, 1930, pp. 305–331.

MOORE, R. L. [1] *Foundations of point-set theory*, American Mathematical Society Colloquium publications, vol. 13, 1932.

——— [2] *Concerning upper semi-continuous collections*, Monatshefte für Mathematik und Physik, vol. 36, 1929, pp. 81–88.

MORREY, C. B. [1] *An analytic characterization of surfaces of finite Lebesque area, Part 1*, American Journal of Mathematics, vol. 57, 1935, pp. 692–702.

NEWMAN, M. H. A. [1] *Elements of the topology of plane sets of points*, Cambridge University Press, 1939.

RADO, T. [1] *Sur l'aire des surfaces courbes*, Acta Szeged, vol. 3, 1927, pp. 131–169.

——— [2] *On the semi-continuity of double integrals in parametric form*, Transactions of the American Mathematical Society, vol. 51, 1942, pp. 336–361.

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

AUTOMORPHISM RINGS OF PRIMARY ABELIAN OPERATOR GROUPS[1]

BY REINHOLD BAER

(Revision received March 7, 1942)

The representation of a ring as the ring of all the automorphisms of a primary abelian operator group expresses significant inner properties of this ring, since there exists essentially at most one such representation of a given ring. Thus it is the object of this investigation to expose invariant qualities of these automorphism rings. They are found to be pecularities of their ideal theory (Chapter III); and from these we select a complete set of postulates for the class of all the automorphism rings of primary abelian operator groups (Sections **12** and **13**); thus proving incidentally that the structure of these rings is completely determined by their ideal theory (Section **14**).

The properties of the ideals in these rings reflect so completely the structure of the underlying operator group that we are able to prove—provided the "rank" of this group is at least three—the essential identity of the group of automorphisms of the ring and of the group of projectivities (= biunivoque and monotone increasing transformations) of the system of admissible subgroups of the underlying operator group (Chapter II).

Two [extreme] special cases may serve as an illustration for these theorems. Every projective geometry of finite dimension not less than three may be represented as the set of admissible subgroups of a suitable primary abelian operator group; and our characterization of the automorphisms of the automophism ring specializes in this case to the (well known) theorem that each automorphism of the automorphism ring (= ring of square matrices with coefficients from a suitable (not necessarily commutative) field) is induced by a so-called semi-linear transformation.

If secondly G is an ordinary primary abelian group with the property that the least common multiple of the orders of the elements in G is a prime power p^m and that G contains at least three independent elements of maximum order, then our theorem states that every automorphism of the automorphism ring is an inner auomorphism.

CHAPTER I: THE RELATIONS OF A GROUP TO ITS AUTOMORPHISM RING

1. The Admissible Subgroups

In this section we collect a number of definitions and facts concerning primary abelian operator groups which will be needed in the future.[2]

[1] Presented to the American Math. Soc. September, 1941.

[2] Proofs and details may be found in Baer (4). For references cp. the Bibliography at the end of the paper.

1A. *Cycles.* The element c in the partially ordered set D is termed a *cycle of order* $n = n(c)$, if exactly $n + 1$ elements in D are parts of c, and if the parts of c form an ordered set; if c is a cycle of order n, then it is possible to write down the system of all its parts in the following fashion: $0 = c^{(n)} < c^{(n-1)} < \cdots < c^{(2)} < c' < c^{(0)} = c$.

All the partially ordered sets considered will be complete lattices satisfying the modular (Dedekind's) law. The smallest element containing a certain set of elements will be termed its join and the greatest element contained in all the elements of some set will be called its meet or cross-cut. The following fact[3] is of great importance:

If the element s in the modular lattice D is the join of cycles whose orders do not exceed n, and if c is a subcycle of s, then the order of c does not exceed n.

The subset S of the complete modular lattice D is said to be *independent*, if the meet of the element s in S and of the join of the other elements in S is 0 for every element s in S. We shall usually impose the further restriction that the elements in an independent set should all be different from 0.

1B. *Rings of Subgroups.* Suppose that G is an abelian group and that the composition of its elements is written as addition. Then a system D of subgroups of G shall be termed a ring of subgroups, if it contains 0 and G, and if it contains the cross-cuts and sums[4] of all the subsets of D. It is well known that every ring of subgroups of G is a complete modular lattice; and it is obvious that a group G may contain sets of subgroups which are complete modular lattices, but not rings of subgroups.

1C. *Operator Groups.* Suppose that G is an abelian group and that E is a ring which contains an identity element 1. We say that G is an operator group over E, or that G admits the elements in E as operators, if there exists to every element g in G and to every element e in E an element ge in G such that $(g' \pm g'')e = g'e \pm g''e$, $g(e' \pm e'') = ge' \pm ge''$, $g(e'e'') = (ge')e'' = ge'e''$, $g1 = g$ for g, g', g'' in G, e, e', e'' in E. A subset S of G is an *E-admissible subgroup* if, and only if, $S = S + S$ and $S = SE$. If x is any element in G, then xE is the *E-admissible subgroup generated by* x.

The set of all the E-admissible subgroups in G is denoted by $D(G;E)$. Clearly $D(G;E)$ is a ring of subgroups of G. The set S of elements in G is said to be *independent*, if the elements in S generate different subgroups which are all different from 0, and if the subgroups generated by the elements in S form an independent set of elements in $D(G;E)$.

1D. *Primary Rings.* The ring E is said to be a primary ring, if it contains an identity element 1, and if there exists in E a two-sided ideal $P = P(E)$ such that

---

[3] Baer (4), Theorem I.2.1.

[4] Cross-cut signifies the set-theoretical cross-cut of the subgroups whereas sum is the group-theoretical sum of the subgroups.

$P^m = 0$ for some (smallest) positive $m = m(E)$ and such that every right- and every left-ideal in E is a power of P. Then there exist elements p in E such that $P = pE = Ep$ and such elements p we term *primes in* E. If p is a prime in E, then every ideal in E has the form $P^i = p^iE = Ep^i$. An element in E does not belong to P if, and only if, it possesses an inverse in E. We note finally that the four properties: $P = 0$, $p = 0$, $m = 1$, and "E is a field" are equivalent.[5]

1E. *Primary Abelian Operator Groups.* If the abelian group G admits the elements in the primary ring E as operators, then G is said to be a *primary abelian operator group* (*over* E). If x is an element in the primary abelian operator group G over E, then xE is a cycle in $D(G;E)$ and every cycle in $D(G;E)$ has this form. The order $n(x)$ of the element x is defined as the order $n(xE)$ of the cycle xE in $D(G;E)$. The order of the element x in G does not exceed n if, and only if, $xP^n = 0$. Mapping the element e in E upon the element xe in xE constitutes an isomorphism of $E/P^{n(x)}$ upon xE which maps ideals upon admissible subgroups. If Z is a cycle in $D(G;E)$, and if $0 \leqq i \leqq n(Z)$, then ZP^i is the uniquely determined subcycle of order $n(Z) - i$ of Z.[6]

Every admissible subgroup of the primary abelian operator group G is a direct sum of cycles (= is the sum of an independent set of cycles = possesses a basis).[6a] If U and V are subgroups in $D(G;E)$, and if $U \leqq V$, then the following condition is necessary and sufficient for U to be a direct summand of G: if the subcycle $Z \neq 0$ of U is contained in a subcycle of order n of V, then there exists a cycle of order n between Z and U.[7]

1F. *Criterion for Primarity.*[8] Suppose that the ring E contains an identity element 1, that E is a cycle[9] of order m in the partially ordered set of all the right-ideals in E, that the abelian group G admits the elements in E as operators, and that[10] $D(G;E)$ contains at least two independent cycles of order m. Then E is primary if, and only if, the following property is satisfied by $D(G;E)$:

If U and V are E-admissible subgroups of G, if $U \leqq V$, and if there exist at most two different subcycles of order 1 of V/U (in $D(G/U;E)$), then (and only then) is V/U a cycle in $D(G/U;E)$.

1G. *Projectivities* are biunivoque and monotone increasing maps of one partially ordered set upon another partially ordered set. If G is a primary abelian

---

[5] Baer (4), (i) to (v) of Section II.2, p. 304.

[6] Baer (r), Theorem II.2.1.

[6a] Baer (4), Theorem A. of Section II.2, p. 310.

[7] The most important special case of this theorem is the fact that every subcycle of maximum order is a direct summand; cp. Baer (4), Theorem I.3.6, II.2.4 and the remarks on p. 310.

[8] Baer (4), Theorem II.2.4.

[9] This hypothesis is readily seen to imply that every right-ideal in E is a two-sided ideal.

[10] This hypothesis is only needed to show the sufficiency of the condition.

operator group over E, if $D(G;E)$ contains at least three independent cycles of order $m(E)$, and if φ is a projectivity of $D(G;E)$ upon $D(G';E')$ for G'an abelian group admitting the elements in the ring E' as operators, then there exists[11] an isomorphism η of G upon G' and an isomorphism η of E upon E' such that $S^\eta = S^\varphi$ for S in $D(G;E)$ and $(ge)^\eta = g^\eta e^\eta$ for g in G and e in E.

2. The Annulets

Suppose that the ring E contains an identity element 1, and that the abelian group G admits the elements in E as operators. Then the transformation φ shall be termed an *automorphism of* G (more accurately: *an* E*-automorphism*), if it is a single-valued function of the elements in G with values in G satisfying $(u \pm v)^\varphi = u^\varphi \pm v^\varphi$, $(ue)^\varphi = u^\varphi e$ for u, v in G and e in E. Defining addition and multiplication of automorphisms in the usual fashion it is readily seen that the automorphisms of G form a ring:[12] *the automorphism ring* $\Lambda = \Lambda(G) = \Lambda(G, E)$ *of the group* G *over* E.

It is the main object of this section to exhibit a class of left-ideals and a class of right-ideals in Λ which each reflect faithfully all the properties of the set $D(G;E)$ of the E-admissible subgroups of G.

If Σ is any subset of Λ, then the *left-annulet* $\Lambda(\Sigma)$ determined by Σ consists of all the automorphisms ξ in Λ satisfying $\xi\Sigma = 0$; and the *right-annulet* $\mathrm{P}(\Sigma)$ consists of all the automorphisms χ in Λ such that $\Sigma\chi = 0$. If T is a subset of G, then the *left-annulet* $\Lambda(T)$determined by T consists of all the automorphisms of G which map G into part of T; and the *right-annulet* $\mathrm{P}(T)$ determined by T consists of all those automorphisms of G which map T upon 0.[13] We observe that left-annulets are left-ideals in Λ and that right-annulets are right-ideals in Λ. The left- and right-annulet determined by the same E-admissible subgroup T of G are connected by the following fundamental relation:[14]

$$\Lambda(T)\mathrm{P}(T) = 0.$$

If $N(\Sigma)$ is the set of all the elements in G which are mapped upon 0 by all the automorphisms in the subset Σ of Λ, then it is readily seen that $N(\Sigma)$ is an E-admissible subgroup of G; the set G^Σ, however, of all the elements g^σ for g in G and σ in Σ need not be a subgroup.

THEOREM 2.1[15]: $\mathrm{P}(\Sigma) = \mathrm{P}(G^\Sigma)$ *and* $\Lambda(\Sigma) = \Lambda(N(\Sigma))$ *for every subset* Σ *of* Λ.

PROOF: The automorphism ξ of G belongs to $\mathrm{P}(\Sigma)$ if, and only if, $\Sigma\xi = 0$; and this is the case if, and only if, $0 = G^{\Sigma\xi} = (G^\Sigma)^\xi$, i.e. if, and only if, ξ is in $\mathrm{P}(G^\Sigma)$; and this proves the first equality. Likewise ξ belongs to $\Lambda(\Sigma)$ if, and

---

[11] Baer (4), Theorem II.3.1.

[12] Cp. Baer (1), Fitting (1), Shoda (1).

[13] The definitions of $\mathrm{P}(T)$ and $\Lambda(T)$ are due to Shoda (1).

[14] The fundamental importance of this relation has been discovered by Shiffman (1) who based his discussion of the fully characteristic subgroups of primary abelian groups on the consideration of pairs of two-sided ideals defined by this relation.

[15] This is the generalization of a theorem proved by Shiffman (1) for two-sided ideals Σ.

only if, $\xi\Sigma = 0$; or equivalent: if, and only if, $G^{\xi\Sigma} = 0$; and this happens if, and only if, $G\xi \leqq N(\Sigma)$ showing the validity of the second equality.

THEOREM 2.2:[16] *If G is a primary abelian operator group over the (primary) ring E, and if S is an E-admissible subgroup of G, then*

$$G^{\Lambda(S)} = S = N(\mathrm{P}(S)) \text{ and} \tag{1}$$

$$\Lambda(S) = \Lambda(\mathrm{P}(S)), \qquad \mathrm{P}(S) = \mathrm{P}(\Lambda(S)). \tag{2}$$

PROOF: It is an immediate consequence of the definition of $\Lambda(S)$ that $G^{\Lambda(S)} \leqq S$. To prove the desired equality we consider an element v of maximum order in G. There exists by **1E** an E-admissible subgroup V of G such that G is the direct sum of V and of the cycle vE. If s is any element in S, then there exists one and only one automorphism φ of G which maps V upon 0 and v upon s, since $n(s) \leqq n(v)$, and since sE and $E/P^{n(s)}$, vE and $E/P^{n(v)}$ are isomorphic. Clearly φ belongs to $\Lambda(S)$. Hence s is an element in $G^{\Lambda(S)}$ so that $S = G^{\Lambda(S)}$.

It follows from the definition of $\mathrm{P}(S)$ that S is mapped upon 0 by $\mathrm{P}(S)$ and that therefore $S \leqq N(\mathrm{P}(S))$. To prove the desired equality let t be any element in G not in S, and W a greatest E-admissible subgroup of G which contains S, but not t. Every E-admissible subgroup of G which contains W as a proper subgroup contains therefore the element t so that the element $W + t$ generates the only subcycle of order 1 of G/W. Applying the fundamental property mentioned under **1F** we deduce that G/W is a cycle whose order certainly does not exceed the maximum order of the elements in G. Thus there exists in G an element h of order $n(G/W)$; and there exists one and only one E-automorphism γ of G which maps W upon 0 and some generating coset of G/W upon h. Clearly γ belongs to $\mathrm{P}(S)$, since $S \leqq W$; and t does not belong to $N(\mathrm{P}(S))$, since t^γ is, exactly as $W + t$, of order 1. This completes the proof of (1).

From (1) and Theorem 2.1 we deduce that $\mathrm{P}(\Lambda(S)) = \mathrm{P}(G^{\Lambda(S)}) = \mathrm{P}(S)$ and that $\Lambda(\mathrm{P}(S)) = \Lambda(N(\mathrm{P}(S))) = \Lambda(S)$.

It is an immediate consequence of Theorem 2.1 and Theorem 2.2, that we need not distinguish between the annulets determined by subsets of G and the annulets which originate from subsets of Λ.

If we define as *dualities* the biunivoque and monotone decreasing maps of one partially ordered set upon another partially ordered set, then we may restate the results obtained in this section in the form of the following *fundamental theorem*:

If G is a primary abelian operator group over the (primary) ring E, then a projectivity of $D(G;E)$ upon the partially ordered set of all the left-annulets in $\Lambda(G;E)$ is defined by mapping the E-admissible subgroups S of G upon the left-annulet $\Lambda(S)$; and a duality of $D(G;E)$ upon the partially ordered set of all the right-annulets in $\Lambda(G;E)$ is defined by mapping the E-admissible subgroup S of G upon the right-annulet $\mathrm{P}(S)$.

REMARK: If G is an abelian group which admits as operators the elements in

[16] This is the generalization of a theorem proved by Shiffman (1) for the special case of fully characteristic subgroups S.

the ring E, if the ring E meets all the requirements for primary rings (**1D**) with the exception of the postulate that a power of P be 0 though every element in G is annihilated by some power of P, then at least one of the following three conditions is satisfied by the E-admissible subgroups of G:

(A) The orders of the cycles in $D(G;E)$ are bounded.

(B) There exists an E-admissible direct summand of G which contains one and only one subcycle of order n, for every positive integer n.

(C) To every positive integer n there exists a direct summand of G which is a cycle in $D(G;E)$ whose order exceeds n.

If (A) is satisfied by $D(G;E)$, then all the results of this section are obviously still valid. If (B) is satisfied, then one may show by an argument similar to the one we used in the proof of Theorem 2.2 that $S = N(\mathrm{P}(S))$ for every S in $D(G;E)$, though S need not equal $G^{\mathrm{A}(S)}$; and if (C) is satisfied by $D(G;E)$, then $S = G^{\mathrm{A}(S)}$, though S and $N(\mathrm{P}(S))$ may be different.

Chapter II: The Automorphisms of the Automorphism Ring

3. Projectivities of the Group and Automorphisms of the Ring

The main object of this section is the proof of the theorem[17] that

The group of projectivities of the partially ordered set $D(G;E)$ of all the E-admissible subgroups of the primary abelian operator group G over the primary ring E and the group of all the automorphisms of the ring $\mathrm{A}(G;E)$ are essentially the same, provided there exist at least three independent cycles of order[18] *$m(E)$ in $D(G;E)$.*

In order to construct the isomorphism linking these two groups to each other we have to consider the *D-automorphisms of G*: these are biunivoque correspondences between the elements in G which preserve addition and which map E-admissible subgroups upon E-admissible subgroups; the D-automorphisms are therefore proper automorphisms of G which induce projectivities of $D(G;E)$ upon itself. If $D(G;E)$ contains at least two independent cycles of order $m(E)$, then a necessary and sufficient condition[19] for the proper automorphism φ of G to be a D-automorphism is the existence of an automorphism φ of the ring E such that $(ge)^{\varphi} = g^{\varphi}e^{\varphi}$ for g in G and e in E. It is readily deduced from this criterion that $\varphi^{-1}\gamma\varphi$ is an E-automorphism of G, whenever φ is a D-automorphism and γ an E-automorphism of G. This shows that each D-automorphism φ induces a projectivity: $S \to S^{\Phi}$ of $D(G;E)$ and an automorphism $\gamma \to \varphi^{-1}\gamma\varphi$ of $\mathrm{A}(G;E)$, if there exist at least two independent cycles of order $m(E)$ in $D(G;E)$ and we shall have attained our object as soon as we have shown that the correlation thus established between projectivities of $D(G;E)$ and automorphisms of $\mathrm{A}(G;E)$ is actually an isomorphism between the two groups, provided there exist at least three independent cycles of order $m(E)$ in $D(G;E)$.

If G is an abelian group, then an *endomorphism of G* is a single-valued function

---

[17] The implications of this theorem as well as some of the relations to known theorems are discussed in section **4**.

[18] For the definition of m(E) cf. **1D**.

[19] This may be derived easily from Baer (4), Theorem II.2.2. or Theorem II.3.1.

of the elements in G which maps elements in G upon elements in G and which preserves addition.

THEOREM 3.1: *If G is a primary abelian operator group over the (primary) ring E, if $D(G;E)$ contains at least two independent*[20] *cycles of order $m(E)$, and if φ is an endomorphism of G, then each of the following properties of φ implies all the others:*

(1) *There exists an element e in E such that $g^\varphi = ge$ for every g.*

(2) *$S^\varphi \leqq S$ for every E-admissible subgroup S of G.*

(3) *$\varphi\alpha = \alpha\varphi$ for every E-automorphism α of G.*

PROOF: It is obvious that (2) and (3) are both consequences of (1). Suppose now that (3) is satisfied by φ and that the E-admissible subgroup S of G is a direct summand of G. Then there exists an E-automorphism γ of G which maps G upon S and which leaves invariant every element in S. If s is any element in S, then $s^\varphi = s^{\gamma\varphi} = s^{\varphi\gamma}$ is an element in S so that both (3) and (2) imply the following condition:

(2′) *$S^\varphi \leqq S$ for every E-admissible direct summand S of G.*

Suppose now that (2′) is satisfied by φ. If b is an element of maximum order in G, then it follows from **1E** that the cycle bE of order $m(E)$ is a direct summand of G; and hence there exists one (and only one) element $e(b)$ in E such that $b^\varphi = be(b)$. If b' and b'' are two independent elements of maximum order, then $b' + b''$ is an element of maximum order; and it follows from $b'e(b') + b''e(b'') = b'^\varphi + b''^\varphi = (b' + b'')^\varphi = (b' + b'')e(b' + b'')$ that $e(b') = e(b' + b'') = e(b'')$. By hypothesis there exist two independent elements v and w of maximum order in G. If b is any element of maximum order in G, then b is independent of at least one[21] of the elements v and w; and we deduce from the fact just derived that $e(b) = e(v) = e(w) = e$ for every element b of maximum order in G. If $g \neq 0$ is any element in G, then g is independent[22] of at least one of the elements v and w. If g and v are independent, then $v + g$ is an element of maximum order so that $ve + g^\varphi = v^\varphi + g^\varphi = (v + g)^\varphi = (v + g)e = ve + ge$ or $g^\varphi = ge$ for every g. Thus (1) is a consequence of (2′), completing the proof of the equivalence of (1), (2) and (3).

COROLLARY 3.2: *Suppose that G is a primary abelian operator group over the (primary) ring E and that $D(G;E)$ contains at least two independent cycles of order $m(E)$.*

(a) *Two D-automorphisms of G induce the same projectivity in $D(G;E)$ if, and only if, they induce the same automorphism in the ring $\mathrm{A}(G;E)$.*

(b) *Mapping the projectivity induced by the D-automorphism φ in $D(G;E)$ upon the automorphism induced by φ in $\mathrm{A}(G;E)$ constitutes an isomorphism of a group*[23] *of projectivities of D upon a group*[23] *of automorphisms of A.*

---

[20] It is readily seen that this hypothesis cannot be omitted.

[21] Since cycles different from 0 contain one and only one cycle of order 1, i.e.contain one and only one smallest admissible subgroup different from 0.

[22] Since gE is a cycle in $D(G;E)$.

[23] Which may or may not be the whole group.

The statement (a) is easily deduced from the equivalence of the properties (2) and (3) of Theorem 3.1; and the statement (b) is an immediate consequence of (a).

LEMMA 3.3: *If G is a primary abelian operator group over the (primary) ring E, if $D(G;E)$ contains two independent cycles of order $m(E)$, and if the automorphism φ of the ring $\mathrm{A}(G;E)$ leaves all the left-annulets in* A *invariant, then* $\varphi = 1$.

PROOF: It is an immediate consequence of Theorem 2.2 that φ leaves all the right-annulets in A invariant; and that the E-automorphism α of G maps therefore the element g upon 0 if, and only if, α^φ maps g upon 0; and α maps G upon the E-admissible subgroup S of G if, and only if, α^φ maps G upon S.

There exists, by **1E**, a basis B of G. To every element b in B and to every element g in G with $n(g) \leqq n(b)$ there exists one and only one E-automorphism $\alpha(b,g)$ of G which maps b upon g and which maps upon 0 all the other elements in B. Then it follows from the remarks in the preceding paragraph of the proof that $\alpha(b, g)^\varphi$ maps the element b upon a generator of gE hence upon an element $ge(b, g)$ for $e(b, g)$ in E, but not in[24] P, and maps all the other elements in B upon 0.

From $\alpha(b, b) = \alpha(b, b)^2$ we infer $\alpha(b, b)^\varphi = \alpha(b, b)^\varphi$ and, by **1.D**, $e(b, b) \equiv e(b, b)^2$ or $e(b, b) \equiv 1 \bmod P^{n(b)}$, i.e. $\alpha(b, b) = \alpha(b, b)^\varphi$.

If b and g are independent elements, then we infer from $\alpha(b, b + g) = \alpha(b, b) + \alpha(b, g)$ that $be(b, b + g) + ge(b, b + g) = (b + g)e(b, b + g) = b^{\alpha(b,b+g)^\varphi} = b^{\alpha(b,b)^\varphi + \alpha(b,g)^\varphi} = b + ge(b, g)$ and that therefore $e(b, b + g) \equiv 1 \bmod P^{n(b)}$ and $e(b, b + g) \equiv e(b, g) \bmod P^{n(g)}$. Hence $e(b, g) \equiv 1 \bmod P^{n(g)}$ and $\alpha(b, g)^\varphi = \alpha(b, g)$.

To every element b in B there exists an element w in G which is independent of b and has the same order as b. If r is any element in E, then b and $g = br + w$ as well as b and w are independent. Hence we may deduce from the result of the preceding paragraph that $\alpha(b, br)^\varphi = \alpha(b, br)^\varphi + \alpha(b, w) - \alpha(b, w) = \alpha(b, br)^\varphi + \alpha(b, w)^\varphi - \alpha(b, w) = \alpha(b, br + w)^\varphi - \alpha(b, w) = \alpha(b, br + w) - \alpha(b, w) = \alpha(b, br)$.

To every element g in G whose order does not exceed $n(b)$ there exists an element r in E and an element w in G whose order does not exceed $n(b)$ and which is either 0 or independent of b such that $g = br + w$. Hence $\alpha(b, g)^\varphi = \alpha(b, br + w)^\varphi = \alpha(b, br)^\varphi + \alpha(b, w)^\varphi = \alpha(b, br) + \alpha(b, w) = \alpha(b, g)$.

If δ is any E-automorphism of G, then $\delta - \alpha(b, b^\delta)$ maps the basis element b upon 0. It follows therefore from a result stated in the first paragraph of this proof that $\delta^\varphi - \alpha(b, b^\delta) = \delta^\varphi - \alpha(b, b^\delta)^\varphi = (\delta - \alpha(b, b^\delta))^\varphi$ maps b upon 0; and this shows that both δ and δ^φ map the basis element b upon b^φ; and this implies the equality of δ and δ^φ, since they are both E-automorphisms of G. This completes the proof of the fact that $\varphi = 1$.

THEOREM 3.4: *If G_i is, for $i = 1, 2$, a primary abelian operator group over the*

[24] For the definition of the ideal P in E, cp. **1D**.

(primary) ring E_i, if $D_i = D(G_i ; E_i)$ contains at least three independent cycles of order $m(E_i)$, then there exists to every isomorphism φ of the ring $\mathfrak{A}_1 = \mathfrak{A}(G_1 ; E_1)$ upon the ring $\mathfrak{A}_2 = \mathfrak{A}(G_2 ; E_2)$ an isomorphism η of G_1 upon G_2 and an isomorphism η of E_1 upon E_2 such that $(ge)^\eta = g^\eta e^\eta$ for g in G_1 and e in E_1 and such that $\alpha^\varphi = \eta^{-1}\alpha^\eta$ for every α in $\mathfrak{A}_1$.

REMARK: A special case of this theorem is the fact that a ring $\mathfrak{A}$ is the ring of E-automorphisms of essentially at most one operator group G meeting our requirements.[25]

PROOF: It is a consequence of Theorem 2.2 that mapping the left-annulet $\mathfrak{A}$ in $\mathfrak{A}_1$ upon $\mathfrak{A}^\varphi$ constitutes a projectivity of the partially ordered set of all the left-annulets in $\mathfrak{A}_1$ upon the partially ordered set of all the left-annulets in $\mathfrak{A}_2$. If S is an E_1-admissible subgroup of G_1, then $S^{\varphi *} = G_2^{\mathfrak{A}(S)^\varphi}$ is by Theorem 2.2, an E_2-admissible subgroup of G_2; and it follows from the theorems of section **2** that φ^* is a projectivity of D_1 upon D_2. There exists by the theorem[26] in **1G** an isomorphism η of G_1 upon G_2 and an isomorphism η of E_1 upon E_2 such that $(ge)^\eta = g^\eta e^\eta$ for g in G_1, e in E_1 and $S^\eta = S^{\varphi *}$ for S in D_1; and we deduce from Theorem 2.2 that

$$\eta^{-1}\mathfrak{A}(S)\eta = \mathfrak{A}(S^\eta) = \mathfrak{A}(S^{\varphi *}) = \mathfrak{A}(G_2^{\mathfrak{A}(S)^\varphi}) = \mathfrak{A}(S)^\varphi \quad \text{for } S \text{ in } D_1 .$$

Applying Lemma 3.3 we find now that $\alpha^\varphi = \eta^{-1}\alpha\eta$ for α in $\mathfrak{A}_1$, since the two isomorphisms under consideration effect the same projectivity in the partially ordered set of the left-annulets in A_1.

COROLLARY 3.5: *If G is a primary abelian operator group over the primary ring E, and if $D(G;E)$ contains at least three independent cycles of order $m(E)$, then every automorphism of $\mathfrak{A}(G;E)$ and every projectivity of $D(G;E)$ upon itself is induced by a D-automorphism of G.*

This is an immediate consequence of Theorem 3.4 and of the Theorem in **1G**. The existence of the desired isomorphism of the group of projectivities of D and the group of automorphisms of $\mathfrak{A}$ is now assured by Corollaries 3.2 and 3.5. This completes the proof of the theorem enunciated at the beginning of this section.

Since the left-ideals of the ring H_n of all the n by n matrices with integral coefficients form a projection of the system of subgroups of a direct sum of n infinite cyclic groups, and since every projection of this latter system upon itself may be induced by an isomorphism of the underlying group,[27] one may prove by a method similar to, though much cruder than the one used in our proof of Theorem 3.4 that *every automorphism of the ring H_n is an inner automorphism.*

[25] Cp. Asano (1), Satz 12, p. 245.

[26] It has been pointed out (Baer (2)) that this theorem would not be true without the hypothesis of the existence of three independent cycles of maximum order. In the proof of Lemma 3.3 we made use of the existence of at least two independent cycles of maximum order. Whether or not either of these hypotheses is needed in order that every automorphism of $\mathfrak{A}(G;E)$ is induced by a $D(G;E)$-automorphism of G, is still an open question.

[27] Baer (2), Theorem 13.1, p. 39.

4. Endomorphisms of the Group and Automorphisms of the Ring

If G is an abelian group, then we denote by $\mathrm{H} = \mathrm{H}(G)$ *the ring of all the endomorphisms of* G. If the group G admits the elements in the ring E as operators, then the ring $\mathrm{A}(G;E)$ of the E-automorphisms of G is a subring of $\mathrm{H}(G)$; and if e is any element in E, then an endomorphism of G is defined by mapping the element g in G upon ge. If the ring E is primary, and if $D(G;E)$ contains cycles of order $m(E)$, then $Ge = 0$ implies $e = 0$, since $te = 0$ implies, by **1E**, $e = 0$ for t an element of order $m(E)$; and hence we may identify the elements in E with the endomorphisms they induce in G so that E may be considered another subring of H. If E is a primary ring, and if $D(G;E)$ contains at least two independent cycles of order $m(E)$, then it follows from Theorem 3.1 that the subrings E and A of H are each the centralizer of the other one in H.[28] This implies in particular that an element in H which possesses an inverse in H transforms E into itself if, and only if, it transforms A into itself; and this fact may be stated as follows: *the subrings* A *and* E *of* H *have the same normalizing group* N *in* H.

THEOREM 4.1: *If* G *is a primary abelian operator group over the (primary) ring* E, *and if* $D(G;E)$ *contains at least three independent cycles of order* $m(E)$, *then the normalizing group* N *of* E *and* A *in* H *consists exactly of the* D*-automorphisms of* G *and all the automorphisms of* E *and of* A *are induced by elements in* N.

PROOF: It has been pointed out at the beginning of section 3, that the D-automorphisms of G transform the ring A into itself, i.e. they belong to N. If conversely ϑ belongs to N, then there exists an inverse endomorphism to the endomorphism ϑ so that ϑ is a proper automorphism of G; and it is a consequence of Corollary 3.5 that there exists a D-automorphism ρ of G satisfying: $\rho^{-1}\alpha\rho = \vartheta^{-1}\alpha\vartheta$ for every α in A. Hence it follows from Theorem 3.1 that $\vartheta\rho^{-1}$ belongs to E and induces the identical projectivity in $D(G;E)$. Thus ρ and $\vartheta\rho^{-1}$ are D-automorphisms of G and consequently every ϑ in N is a D-automorphism. It is a consequence of Corollary 3.5 that every automorphism of A is induced by D-automorphisms of G; and it is well known[29] that every automorphism of E is induced by a D-automorphism of G. Hence every automorphism of the rings E and A is induced by elements in N.

LEMMA 4.2: *Suppose that* U *and* V *are subrings of the ring* W, *that* U *is the centralizer of* V *and* V *the centralizer of* U *in* W, *and that the element* r *in* W *possesses an inverse in* W. *Then* r *induces in* V *an inner automorphism of* V *if, and only if, it induces in* U *an inner automorphism of* U.

PROOF: If r induces in U an inner automorphism of U, then there exists an element s in U which possesses an inverse in U and which satisfies $s^{-1}us = r^{-1}ur$ for every u in U. The element $t = rs^{-1}$ belongs to the centralizer of U so that t is an element in V which possesses an inverse in V. Since s belongs to the centralizer U of V, we find that $t^{-1}vt = r^{-1}vr$ for every v in V, i.e. r induces in V an inner automorphism of V.

---

[28] This implies in particular the equality of the centrum of $\mathrm{A}(G;E)$ and of the centrum of the ring E, a fact that has already been noted by K. Asano (1), Satz 7, p. 242.

[29] This is an immediate consequence of the existence of a basis of G; cp. **1E**.

THEOREM 4.3: *Suppose that G is a primary abelian operator group over the (primary) ring E, and that $D(G;E)$ contains at least three independent cycles of order $m(E)$. Then every automorphism of the ring $\mathfrak{A}(G;E)$ is an inner automorphism of $\mathfrak{A}$ if, and only if, every automorphism of the ring E is an inner automorphism of E.*

This is an immediate consequence of Theorem 4.1 and Lemma 4.3, and of the fact that E and $\mathfrak{A}$ are each the centralizer of the other in H.

We mention two interesting special cases:

1. If G is an ordinary primary abelian group the orders of whose elements are bounded and which contains at least three independent elements of maximum order, then every automorphism of the ring $\mathfrak{A}$ of automorphisms of G is an inner automorphism of $\mathfrak{A}$.

For E is the ring of integers modulo a power of a prime.

2. If n is a finite or infinite cardinal number not less than 3, and if $\mathfrak{A}$ is the ring of all the row-finite, n by n matrices with coefficients from the field of real numbers (or the field of quaternions or some other field all of whose automorphisms are inner), then every automorphism of $\mathfrak{A}$ is an inner automorphism.

For we may choose for G the group of rank n over the field of coefficients.

Every inner automorphism of a ring is a center automorphism[30] of this ring; and it is a well known theorem[31] that all center automorphisms of a finite simple algebra are inner automorphisms. Finite simple algebras are known[32] to be rings of square matrices with coefficients from a field which is finite over its center; and thus they are automorphism rings of primary abelian operator groups. The following example shows that automorphism rings of primary abelian operator groups may possess center automorphisms which are not inner automorphisms in spite of finiteness hypotheses.

It is a consequence of Lemma 4.2 that it suffices to construct a *primary ring E which is finite over its center and which possesses outer center automorphisms.*

Denote by F a commutative field whose characteristic is not 2 and by Q the field $F(x)$ of all the rational functions in an indeterminate x with coefficients in F. There exists one and only one automorphism φ of Q which maps x upon $-x(\neq x)$ and which leaves invariant all the elements in F. If q is an element not 0 in Q, then $q^{\varphi}q^{-1}$ is readily verified to be an element different from x and x^{-1}.

Denote by E the set of all the ordered pairs (a, b) of elements in Q subject to the following rules: $(a, b) = (c, d)$ if, and only if, $a = c$ and $b = d$; $(a, b) \pm (c, d) = (a \pm c, b \pm d)$, $(a, b)(c, d) = (ac, bc^{\varphi} + ad)$. It is easily seen that E is a ring with identity $1 = (1, 0)$, that the element (a, b) in E possesses an inverse in E if, and only if, $a \neq 0$; and that the only ideal (right- or left-) in E which is different from 0 and E is the two-sided ideal P consisting of all the elements $(0, b)$. Thus E is a primary ring whose center consists exactly of the elements $(a, 0)$ for a in $F(x^2)$, showing that the rank of E over its center is exactly 4. An

[30] i.e. an automorphism which leaves invariant every element in the center of this ring.

[31] Brauer (1), Skolem (1).

[32] Cp. e.g. Albert (1), Theorem 9, p. 39.

automorphism ψ of E is defined by $(a, b)^{\psi} = (a, bx)$ and this automorphism is a center automorphism, since $(a, 0) = (a, 0)^{\psi}$. If (r, s) for $r \neq 0$ were to induce the automorphism ψ in E, then $(0, r) = (r, s)(0, 1) = (0,1)^{\psi}(r, s) = (0, x)(r, s) = (0, xr^{\varphi})$ or $x = r(r^{\varphi})^{-1}$; and the impossibility of this equation has been pointed out before. Thus E meets all the requirements.

Chapter III: Ideal-Theoretical Properties of the Annulets

5. Finite Sums of Cycles and Finite Cross-cuts of Anti-cycles

An element s in the partially ordered set S is termed an *anti-cycle of index n in S*, if the elements in S which contain s form a cycle of order n. The E-admissible subgroup V of the primary abelian operator group G over the (primary) ring E is, by the Fundamental Theorem of section **2**, a cycle (an anticycle) in $D(G;E)$ if, and only if, $\mathfrak{A}(V)$ is a cycle (an anti-cycle) in the partially ordered set of all the left-annulets in $\mathfrak{A}(G;E)$; and this is the case if, and only if, $\mathfrak{P}(V)$ is an anti-cycle (a cycle) in the partially ordered set of all the right-annulets in $\mathfrak{A}(G;E)$.[33]

Lemma 5.1: *If G is a primary abelian operator group over the (primary) ring E, if Z is a cycle and S an anti-cycle in $D(G;E)$, then every E-automorphism of Z is induced by an E-automorphism of G; and every E-automorphism of G/S is induced by an E-automorphism of G which maps S into part of S.*

Proof: If $Z = zE$, and if φ is an E-automorphism of Z, then $z^{\varphi} = ze$ for some e in E and φ is completely determined by e. There exists by **1E** a basis B of G and hence there exist elements $d(b)$ in E almost all of which are 0 such that $z = \sum_{b \text{ in } B} bd(b)$. There exist elements $e(b)$ in E satisfying $d(b)e = e(b)\, d(b)$, since E is a primary ring and its left-ideals are right-ideals. There exists finally one and only one E-automorphism γ of G which maps b upon $be(b)$; and γ induces φ in Z, since γ maps z upon ze.

The basis B contains at least one element, say w, which generates G modulo S, since otherwise all the cosets $S + b$ for b in B would be contained in $(G/S)P$ (by **1E**) so that B would be part of $GP < G$, an impossibility. Since the elements in the cycle wE represent all the cosets of G/S, there exists to every element $b(\neq w)$ in B an element $r(b)$ in E such that $b - wr(b)$ is in S. If p is a prime in E (**1D**) and if n is the order of the cycle G/S, then wp^n is an element in S. The E-admissible subgroup S^* generated by wp^n and the elements $b - wr(b)$ is part of S; and G/S^* is a cycle whose order does not exceed n, since every coset in G/S^* is represented by an element in the cycle wE, and since $wP^n \leqq S^*$. Consequently $S = S^*$.

If $W = S + w$, and if σ is an E-automorphism of G/S, then $W^{\sigma} = We$ for e in E; and σ is completely determined by e. There exist elements $c(b)$ in E such that $er(b) = r(b)c(b)$, since E is a primary ring and its right-ideals are left-ideals; and

[33] If S is an anti-cycle in $D(G;E)$, z an element in G, but not in $S + GP$, then $G = S + zE$ where zE is a cycle in $D(G;E)$ and the index of S in $D(G;E)$ does not exceed the order of zE.

there exists one and only one E-automorphism τ of G which maps w upon we and $b \neq w$ in B upon $bc(b)$. Since $(b - wr(b))^\tau = bc(b) - wer(b) = (b - wr(b))c(b)$, it follows that τ maps the elements $b - wr(b)$, wp^n which generate S upon elements in S so that $S^\tau \leqq S$. The E-automorphism τ maps consequently every coset of G/S into part of a well-determined coset of G/S and τ maps in particular the generating coset W into part of We. Thus we have shown that σ is induced by τ in G/S.

LEMMA 5.2: *If G is a primary abelian operator group over the (primary) ring E, and if G is a cycle of order n in $D(G;E)$, then $\mathfrak{A}(G;E)$ and E/P^n are anti-isomorphic rings.*

PROOF: There exists an element g in G such that $G = gE$. If φ is any E-automorphism of G, then $g^\varphi = ge$ where $P^n + e$ is uniquely determined by g and φ and where g and e determine φ completely. If η is another E-automorphism of G, $g^\eta = gd$, then $g^{\varphi\eta} = (ge)^\eta = g^\eta e = gde$; and now it is obvious how to complete the proof of our statement.

If G is a primary abelian operator group over the (primary) ring E, and if G/S for S an E-admissible subgroup is the sum of a finite number of cycles in $D(G/S;E)$, then it follows from **1E** that G/S is the direct sum of a finite number of cycles and therefore S is the cross-cut of a finite number of anti-cycles. From this fact one deduces easily the equivalence of the following three properties of the E-admissible subgroup S of G: (1) G/S is the sum of a finite number of cycles in $D(G;E)$. (2) S is the cross-cut of a finite number of anti-cycles in $D(G;E)$. (3) $\mathfrak{P}(S)$ is the sum of a finite number of cycles in the partially ordered set of all the right-annulets in $\mathfrak{A}(G;E)$.

LEMMA 5.3: *If G is a primary abelian operator group over the (primary) ring E, if the E-admissible subgroup S of G is the sum of a finite number of cycles in $D(G;E)$, and if the E-admissible subgroup T of G is the cross-cut of a finite number of anti-cycles in $D(G;E)$, then S is part of an E-admissible direct summand of G which is the sum of a finite number of cycles; and T contains an E-admissible direct summand of G which is the cross-cut of a finite number of anti-cycles in $D(G;E)$.*

PROOF: Our statement concerning S is an almost obvious consequence of the fact that there exists a basis of G (**1E**). Denote by W the set of all the E-admissible subgroups H of G which are direct summands of G and which are cross-cuts of a finite number of anti-cycles in G. There exists a subgroup K in W such that $(T + K)/T$ is as small as possible, since the minimum condition is satisfied by the E-admissible subgroups of G/T. If K were not part of T, then $T < T + K$ and there would exist an E-admissible subgroup V between T and $T + K$ such that $(T + K)/V$ is a cycle of order 1. Since K is not part of V, there exists in K an element u of least order which is not in V. Clearly $uP \leqq V$ and $V + uE = T + K$. There exists by **1E** a basis B of K and there exists in B a finite number of elements $b_1, \cdots, b_j$ such that u is in $\sum_{i=1}^{j} b_i E$. Denote by B' the set of those elements in B which are different from all the b_i and which are contained in V; and by B'' the set of the elements in B which are

different from the b_i and which are not in V. If b is in B'', then there exists an element $e(b)$ in E such that $b^* = b - ue(b)$ is in V. Since b is in $T + K$, but not in V, it follows that $n(ue(b)) \leqq n(u) \leqq n(b)$. Since u is independent of the set of elements in B' and B'', it follows that a basis B^* of K is formed by the elements b_i, the elements in B' and the elements b^* for b in B''. It is readily verified that the E-admissible subgroup U of G which is generated by the elements in B^* different from b_i belongs to W, though $T + U \leqq V < T + K$, a contradiction which proves our contention.

6. The Finite Automorphisms

If φ is an E-automorphism of the abelian group G over the primary ring E, then G^φ and $G/N(\varphi)$ are isomorphic groups over the operator ring E; thus G^φ is the sum of a finite number of cycles if, and only if, $N(\varphi)$ is the cross-cut of a finite number of anti-cycles. The E-automorphism φ of G is termed *finite*, if G^φ is the sum of a finite number of cycles and $N(\varphi)$ the cross-cut of a finite number of anti-cycles.

LEMMA 6.1: *If G is a primary abelian operator group over the (primary) ring E, and if φ is a finite E-automorphism of G, then there exist E-admissible subgroups U and V of G with the following properties:*

(i) *G is the direct sum of U and V.*

(ii) *U is the sum of a finite number of cycles in $D(G;E)$.*

(iii) *$G^\varphi \leqq U$ and $V \leqq N(\varphi)$.*

PROOF: There exist by Lemma 5.3 E-admissible subgroups A and B of G such that G is the direct sum of A and B, A is the sum of a finite number of cycles in $D(G;E)$ and $B \leqq N(\varphi)$. Denote by C the cross-cut of B and $A + G^\varphi$. Then C is an E-admissible subgroup of B and C is the sum of a finite number of cycles in $D(B;E)$. Hence there exist by Lemma 5.3 E-admissible subgroups V and W of B such that B is the direct sum of V and W, W is the sum of a finite number of cycles in $D(G;E)$ and $C \leqq W$. Then G is the direct sum of $U = A + W$ and V, U is the sum of a finite number of cycles and $G^\varphi \leqq U$, $V \leqq N(\varphi)$.

We denote by $\Phi = \Phi(G;E)$ *the set of all the finite E-automorphisms of G.*

LEMMA 6.2: *If G is a primary abelian operator group over the (primary) ring E, then $\Phi(G;E)$ is a two-sided, idempotent ideal in the ring $\mathrm{A}(G;E)$; and both the sum of all the left-annulets $\Lambda(Z)$ for cyclic Z and the sum of all the right-annulets $\mathrm{P}(H)$ for anti-cyclic H are equal to $\Phi(G;E)$.*

PROOF: If φ and γ are finite E-automorphisms, and if ρ is an E-automorphism of G, then we infer from $G^{\varphi \pm \gamma} \leqq G^\varphi + G^\gamma$ and $G^{\rho\varphi} \leqq G^\varphi$ that Φ is a left-ideal in A. Since $G^{\varphi\rho}$ is a linear map of G^φ, both subgroups are generated by a finite number of elements, showing that Φ is a two-sided ideal in A. If φ is a finite E-automorphism of G, then there exist by Lemma 6.1 E-admissible subgroups U and V of G meeting the requirements (*i*) to (*iii*)of Lemma 6.1. If ϵ is the E-automorphism of G which maps V upon 0 and leaves invariant every element in U, then ϵ is an idempotent such that $\epsilon\varphi = \varphi\epsilon = \varphi$ and ϵ is finite; and hence Φ is idempotent. It is a consequence of **1E** that $G/N(\varphi)$ is the direct sum of a

finite number of cycles $Z_1, \cdots, Z_k$. There exists one and only one E-automorphism φ_i of G which induces φ in Z_i and which maps the Z_j for $j \neq i$ upon 0. Clearly $\varphi = \varphi_1 + \cdots + \varphi_k$, G^{φ_i} is a cycle and $N(\varphi_i)$ an anti-cycle; and this completes the proof.

We note several consequences of this Lemma 6.2:

1. The identity-automorphism of G is finite if, and only if, G itself is the sum of a finite number of cycles in $D(G;E)$; and Φ is a (right- or left-) annulet if, and only if, $\Phi = \mathfrak{A}$.

2. Every right- (left-) ideal in the subring Φ of $\mathfrak{A}$ is a right- (left-) ideal in the ring $\mathfrak{A}$.

3. If S and T are E-admissible subgroups of G, then each of the following equalities may be easily seen to imply the others:

$S = T$, $\mathfrak{P}(S) \cap \Phi = \mathfrak{P}(T) \cap \Phi$, $\Lambda(S) \cap \Phi = \Lambda(T) \cap \Phi$.

Comparing 2. and 3. we note that the ideal theory in the rings Φ and $\mathfrak{A}$ is different, provided these rings are different; but that the theory of annulets is essentially the same in both rings.

It is known[34] that the E-automorphisms of G may be approximated by means of finite E-automorphisms. It is the object of the following remarks to describe this approximation in terms which involve properties of the ring $\mathfrak{A}$ only. With this in mind we denote by $\Omega = \Omega(G;E)$ *the set of all the right-annulets* $\mathfrak{I}$ *in* $\mathfrak{A}$ *which are cross-cuts of a finite number of anti-cycles in the partially ordered set of the right-annulets in* $\mathfrak{A}$; and it will be convenient to term *Ω-function* any single-valued function $\varphi(\mathfrak{I})$ of the right-annulets in Ω with the following properties:

$\varphi(\mathfrak{I})$ *is a coset of* $\mathfrak{A}/\mathfrak{I}$; *and* $\mathfrak{I} \leqq \mathfrak{I}'$ *implies* $\varphi(\mathfrak{I}) \geqq \varphi(\mathfrak{I}')$.

If φ is any E-automorphism of G, then $\mathfrak{I} + \varphi$ for $\mathfrak{I}$ in Ω is the typical example of such an Ω-function.

THEOREM 6.3: *Suppose that G is a primary abelian operator group over the (primary) ring E.*

(a) *The right-annulet $\mathfrak{I}$ is in Ω if, and only if, $\Lambda(\mathfrak{I}) \leqq \Phi$.*

(b) *To every $\mathfrak{I}$ in Λ and φ in $\mathfrak{A}$ there exists a finite E-automorphism r of G such that $\varphi \equiv \rho \bmod \mathfrak{I}$.*

(c) *To every Ω-function $\varphi(\mathfrak{I})$ there exists one and only one E-automorphism φ of G such that $\varphi(\mathfrak{I}) = \mathfrak{I} + \varphi$ for every $\mathfrak{I}$ in Ω.*

It seems desirable to find an expression of the closure property (*c*) of the ring $\mathfrak{A}$ which does not involve the elements of $\mathfrak{A}$, but which refers to the annulets only.

PROOF: If S is an E-admissible subgroup of G, B a basis of S and C a basis of G, then the (cardinal) number of the elements in B whose order is at least n does not exceed the (cardinal) number of the elements in C whose order is at least n. Hence there exists an E-automorphism φ of G such that $G^{\varphi} = S$; and consequently $\Lambda(\mathfrak{I}) \leqq \Phi$ if, and only if, $G^{\Lambda(\mathfrak{I})}$ is the sum of a finite number of cycles in $D(G;E)$. It is a consequence of Theorems 2.1 and 2.2 that $G^{\Lambda(\mathfrak{I})} = N(\mathfrak{I})$; and

[34] Baer (1).

that $N(\mathrm{I})$ is the sum of a finite number of cycles in $D(G;E)$ if, and only if, $\mathrm{I} = \mathrm{P}(N(\mathrm{I}))$ is the cross-cut of a finite number of anti-cycles in the partially ordered set of the right-annulets in Λ. It is a consequence of Lemma 5.3 that G is the direct sum of E-admissible subgroups U and V such that $N(\mathrm{I}) \leqq U$ and such that U is the sum of a finite number of cycles in $D(G;E)$. Then there exists to every E-automorphism φ of G one and only one E-automorphism γ of G which coincides with φ on U and which maps V upon 0. It is clear that γ is a finite E-automorphism satisfying $\varphi \equiv \gamma \bmod \mathrm{I}$.

We note that $\mathrm{P}(x)$ is, for every x in G, a right-annulet in the class Ω, since $\Lambda(\mathrm{P}(x)) = \Lambda(xE) \leqq \Phi$. If φ and γ are any two E-automorphisms of G satisfying $\mathrm{I} + \varphi = \mathrm{I} + \gamma$ for every I in Ω, then $\varphi - \gamma = 0$ as an element in the cross-cut 0 of all the $\mathrm{P}(x)$ for x in G. If finally $\varphi(\mathrm{I})$ is an Ω-function, then all the elements in the coset $\varphi(\mathrm{P}(x))$ have the same value x^{φ} on the element x in G. If x is in G and e in E, then $xeE \leqq xE$ implies $\mathrm{P}(x) = \mathrm{P}(xE) \leqq \mathrm{P}(xeE) = \mathrm{P}(xe)$ so that $(xe)^{\varphi} = (xe)^{\varphi(\mathrm{P}(x))} = x^{\varphi(\mathrm{P}(x))}\, e = x^{\varphi}e$. If x and y are elements in G, then $\varphi(\mathrm{P}(xE + yE))$ is part of $\varphi(\mathrm{P}(x))$, $\varphi(\mathrm{P}(y))$ and $\varphi(\mathrm{P}(x \pm y))$ so that $(x \pm y)^{\varphi} = (x \pm y)^{\varphi(\mathrm{P}(xE+yE))} = x^{\varphi(\mathrm{P}(x))} + y^{\varphi(\mathrm{P}(y))} = x^{\varphi} \pm y^{\varphi}$. Thus φ is an E-automophism of G. If I is in Ω and x in $N(\mathrm{I})$, then $\mathrm{I} = \mathrm{P}(N(\mathrm{I})) \leqq \mathrm{P}(x)$ and hence $\varphi(\mathrm{I}) \leqq \varphi(\mathrm{P}(x))$; and this implies that $x^{\varphi} = x^{\varphi(\mathrm{I})}$ for x in $N(\mathrm{I})$ or $\varphi(\mathrm{I}) = \mathrm{I} + \varphi$.

7. Sums of Annulets

It is easily seen that the cross-cut of any set of right-(left-)annulets is a right-(left-)annulet. Thus there exists to every set of right-(left-)annulets a smallest right-(left-) annulet, the join of the annulets in the set. But in general this join will be different from the sum of the annulets in the set.

THEOREM 7.1: *Suppose that G is a primary abelian operator group over the (primary) ring E, and that P_i is a right-annulet, Λ_i a left-annulet in $\Lambda(D;E)$.*

(a) *If $\Lambda_1 \cap \Lambda_2 = 0$, then $\Lambda_1 + \Lambda_2$ is a left-annulet.*

(b) *If $\mathrm{P}_1 \cap \mathrm{P}_2 = 0$, then $\mathrm{P}_1 + \mathrm{P}_2$ is a right-annulet.*

(c) *If $\mathrm{P}_1 = \mathrm{P}(\Lambda_i)$, $\Lambda_i = \Lambda(\mathrm{P}_i)$ and $\mathrm{P}_1 \cap \mathrm{P}_2 = 0 = \Lambda_1 \cap \Lambda_2$, then $\mathrm{P}_1 + \mathrm{P}_2 = \Lambda = \Lambda_1 + \Lambda_2$.*

PROOF: It is an immediate consequence of the Fundamental Theorem of section **2** that the cross-cut of $S_1 = G^{\Lambda_1}$ and $S_2 = G^{\Lambda_2}$ is $G^{\Lambda_1 \cap \Lambda_2} = 0$, and that the smallest left-annulet which contains both Λ_1 and Λ_2 is $\Lambda(S_1 + S_2)$. If φ is any element in $\Lambda(S_1 + S_2)$ and g any element in G, then g^{φ} is an element in the direct sum $S_1 + S_2$. Hence there exist uniquely determined elements g_1 and g_2 in S_1 and S_2 respectively such that $g^{\varphi} = g_1 + g_2$. The transformation φ_i which maps g upon g_i is readily seen to be an E-automorphism in $\Lambda_i = \Lambda(S_i)$; and $\varphi = \varphi_1 + \varphi_2$. Hence $\Lambda_1 + \Lambda_2 = \Lambda(S_1 + S_2)$.

If $T_i = N(\mathrm{P}_i)$, then it is an immediate consequence of the Fundamental Theorem of section **2** that $T_1 + T_2 = N(\mathrm{P}_1 \cap \mathrm{P}_2) = G$ and that $\mathrm{P}(T_1 \cap T_2)$ is the smallest right-annulet which contains both P_1 and P_2. Since $G/(T_1 \cap T_2)$ is the direct sum of $T_1/(T_1 \cap T_2)$ and $T_2/(T_1 \cap T_2)$, there exists to every E-automorphism φ in $\mathrm{P}(T_1 \cap T_2)$ one and only one E-automorphism φ_i which maps T_i upon

0 and which coincides with φ on $T_{i\pm1}/(T_1 \cap T_2)$. Clearly $\varphi = \varphi_1 + \varphi_2$ and φ_i is in $\mathfrak{P}_i = \mathfrak{P}(T_i)$. Hence $\mathfrak{P}_1 + \mathfrak{P}_2 = \mathfrak{P}(T_i \cap T_2)$.

If the hypotheses of (c) are satisfied by Λ_i, $\mathfrak{P}_i$, then it follows from what has already been shown that $S_i = G^{\Lambda_i} = N(\mathfrak{P}_i)$, that G is the direct sum of S_1 and S_2, $\Lambda_1 + \Lambda_2 = \Lambda(G) = \Lambda = \mathfrak{P}(0) = \mathfrak{P}_1 + \mathfrak{P}_2$.

If G is an abelian operator group over the ring E, if S is an E-admissible subgroup of G, and if φ is a single-valued additive function of the elements in S with values in G satisfying $(se)^\varphi = s^\varphi e$ for s in S and e in E, then φ is *a linear transformation of S into G.*

LEMMA 7.2: *If the primary abelian operator group G over the (primary) ring E is the direct sum of cycles of equal order m in $D(G;E)$, if S is an E-admissible subgroup of G, and if φ is a linear transformation of S into G, then φ is induced in S by an E-automorphism of G.*

PROOF: If p is a prime in E (cf. **1G**), then every element in G has the form gp^i for g an element of order m in G. Hence there exists a basis B of G and integers $h(b)$ for b in B such that the E-admissible subgroup S of G is generated[35] by the elements $bp^{h(b)}$. Since $n((bp^{h(b)})^\varphi) \leqq n(bp^{h(b)}) = m - h(b)$, there exist elements b' in G such that $b'p^{h(b)} = (bp^{h(b)})^\varphi$. There exists one and only one E-automorphism ψ of G which maps every b in B upon the corresponding b', since $n(b') \leqq n(b) = m$; and ψ clearly induces φ in S.

THEOREM 7.3: *The following three properties of the primary abelian operator group G over the (primary) ring E imply each other:*

(a) *G is the direct sum of cycles of equal order m in $D(G;E)$.*

(b) *The sum of any two right-annulets in $\mathfrak{A}(G;E)$ is a right-annulet.*

(c) *The sum of any two left-annulets in $\mathfrak{A}(G;E)$ is a left-annulet.*

PROOF: Suppose that condition (a) is satisfied by G. It is a consequence of the Fundamental Theorem in section **2** that any two right annulets in $\mathfrak{A}$ have the form $\mathfrak{P}(S)$, $\mathfrak{P}(T)$ where S and T are E-admissible subgroups of G, and that $\mathfrak{P}(S \cap T)$ is the smallest right-annulet containing both $\mathfrak{P}(S)$ and $\mathfrak{P}(T)$. If v is an E-automorphism in $\mathfrak{P}(S \cap T)$, then there exists a linear transformation ψ of $S + T$ into G which coincides with v on S and which maps T upon 0, since $S + T$ is modulo $S \cap T$ the direct sum of S and T, and since $(S \cap T)^v = 0$. There exists by Lemma 7.2 an E-automorphism ω of G which induces ψ in $S + T$. Thus ω belongs to $\mathfrak{P}(T)$ and $v - \omega$ belongs to $\mathfrak{P}(S)$; and hence v belongs to $\mathfrak{P}(S) + \mathfrak{P}(T)$; i.e. (a) implies (b).

It is a consequence of the Fundamental Theorem of section **2** that any pair of left-annulets in $\mathfrak{A}$ may be represented in the form $\Lambda(S)$, $\Lambda(T)$ where S and T are E-admissible subgroups of G, and that $\Lambda(S + T)$ is the smallest left-annulet containing both $\Lambda(S)$ and $\Lambda(T)$. If φ is an E-automorphism in $\Lambda(S + T)$, B a basis of G, then b^φ is an element in $S + T$ and there exist therefore elements $s(b)$, $t(b)$ in S and T respectively such that $b^\varphi = s(b) + t(b)$. Since b is by

[35] Note that the elements $bp^{h(b)}$ which are different from 0 form a basis of S. The existence of a basis of S is assured by **1E** and such a basis would be the starting point for the construction of B.

(a) an element of maximum order m in G, there exist uniquely determined E-automorphisms η and κ such that $b^\eta = s(b)$, $b^\kappa = t(b)$. Clearly $\varphi = \eta + \kappa$ and η is in $\Lambda(S)$, κ in $\Lambda(T)$ so that (c) is a consequence of (a).

If condition (a) is not satisfied by the group G, then G is the direct sum of E-admissible subgroups U, V, W where W may be 0 and where U and V are cycles in $D(G; E)$ such that $0 < n(U) < n(V) =$ the maximum-order of the cycles in $D(G; E)$. There exist elements u and v of order 1 in U and V respectively. The cycles uE and $(u + v)E$ are of order 1 and their cross-cut is 0. Thus $\mathfrak{A}$ is the smallest right-annulet containing both $\mathrm{P}(u)$ and $\mathrm{P}(u + v)$. If the identical automorphism 1 were in $\mathrm{P}(u) + \mathrm{P}(u + v)$, then there would exist E-automorphisms η, κ such that $u^\eta = 0$, $(u + v)^\kappa = 0$ and $\eta + \kappa = 1$; and this would imply: $u = u^{\eta+\kappa} = u^\kappa$, $0 = (u + v)^\kappa = u + v^\kappa$ or $uE = (vE)^\kappa \leqq GP^{n(V)-1}$ which is impossible, since $UP^{n(V)-1} \leqq UP^{n(U)} = 0$. Consequently (b) implies (a). If $V = wE$, then $(u + w)E$ is a cycle of order $n(V)$ whose cross-cut with V is exactly VP. Automorphisms in $\Lambda(V)$ and $\Lambda((u + w)E)$ map elements of an order not exceeding $n(V) - 1$ upon elements in VP; and every automorphism in $\Lambda(V) + \Lambda((u + w)E)$ maps therefore the elements of an order not exceeding $n(V) - 1$ upon elements in VP. But there exist E-automorphisms which map U upon uE, W upon 0 and leave invariant every element in V. Such an automorphism would be in $\Lambda(V + (u + w)E) = \Lambda(V + uE)$, but not in $\Lambda(V) + \Lambda((u + w)E)$, since it maps an element in V of an order less than $n(V)$ upon the element u not in VP; and hence (a) is a consequence of (c).

COROLLARY 7.4: *If the primary abelian operator group G over the (primary) ring E is the direct sum of cycles of equal order m in $D(G; E)$, then*

(i) $\varphi\mathfrak{A} = \mathrm{P}(N(\varphi))$ *and* $\mathfrak{A}\varphi = \Lambda(G^\varphi)$ *for every E-automorphism φ of G;*

(ii) *every right-(left-)annulet is a principal right-(left-)ideal in $\Lambda(G; E)$;*

(iii) *every right-(left-)ideal in $\mathfrak{A}$ which is generated by a finite number of elements*[36] *is a principal right-(left-)ideal in $\mathfrak{A}$.*

PROOF: If γ is in $\mathrm{P}(N(\varphi))$, then $N(\varphi) \leqq N(\gamma)$. To every element x in G^φ there exists one and only one coset $x^{\varphi^{-1}}$ of $G/N(\varphi)$ which is mapped upon x by φ. Thus $N(\gamma) + x^{\varphi^{-1}}$ is a single, well determined coset of $G/N(\gamma)$ and $x^\eta = (N(\gamma) + x^{\varphi^{-1}})^\gamma$ is a uniquely determined element in G^γ. This transformation η is readily seen to be a linear transformation of G^φ into G; and hence it follows from Lemma 7.2 that η is induced in G^φ by an E-automorphism κ of G. If x is any element in G, then $x^{\varphi\kappa} = x^{\varphi\eta} = (N(\gamma) + x)^\gamma = x^\gamma$ or $\varphi\kappa = \gamma$ and this proves $\varphi\mathfrak{A} = \mathrm{P}(N(\varphi))$.

If ρ is an automorphism in $\Lambda(G^\varphi)$, then $G^\rho \leqq G^\varphi$ and the set T of all the elements in G which are mapped by φ upon elements in G^ρ is an E-admissible subgroup such that $N(\varphi) \leqq T$, $T^\varphi = G^\rho$. If Y is any coset of $G/N(\rho)$, then Y^ρ is a well determined element in $G^\rho = T^\varphi$ and there exists one and only one coset $Y^\sigma = Y^{\rho\varphi^{-1}}$ of $T/N(\varphi)$ which is mapped by φ upon the element Y^ρ. It is readily seen that σ is an isomorphism of the group $G/N(\rho)$ over E upon the group $T/N(\varphi)$ over E. There exists a basis B of G. If b is an element in B, then denote by b^*

[36] Has a finite basis in the customary terminology.

some element in $(N(\rho) + b)^\sigma$ so that $(N(\rho) + b)^\sigma = N(\varphi) + b^*$ for every b in B. Since all the elements b in B are of the maximum-order m in G, there exists one and only one E-automorphism τ of G which maps b upon b^* for every b in B. Then $b^{\tau\varphi} = b^{*\varphi} = (N(\varphi) + b^*)^\varphi = (N(\rho) + b)^{\sigma\varphi} = (N(\rho) + b)^\rho = b^\rho$ or $\tau\varphi = \rho$ and this proves $\mathfrak{A}\varphi = \Lambda(G^\varphi)$.

It is a consequence of the Fundamental Theorem of section **2** that every right-(left-)annulet in $\mathfrak{A}$ has the form $\mathrm{P}(S)$ $(\Lambda(S))$ for S an E-admissible subgroup of G; and it is readily verified that there exist to every E-admissible subgroup S of G E-automorphisms φ, γ such that $S = N(\varphi)$ and $S = G^\gamma$. Now (ii) is a consequence of (i). The assertion (iii) finally is an immediate consequence of (i), (ii) and Theorem 7.3.

THEOREM 7.5: *The following five properties of the primary abelian operator group G over the (primary) ring E are equivalent:*

(1) *G is the direct sum of a finite number of cycles of equal order.*[37]

(2) *Sums*[38] *of right-annulets in $\mathfrak{A}(G; E)$ are right-annulets.*

(3) *Every right-ideal in $\mathfrak{A}$ is a right-annulet.*

(4) *Sums*[38] *of left-annulets in $\mathfrak{A}$ are left-annulets.*

(5) *Every left-ideal in $\mathfrak{A}$ is a left-annulet.*[39]

PROOF: It is obvious that (3) implies (2) and that (5) implies (4). If (2) or (4) is satisfied by G, then G is, by Theorem 7.3, the direct sum of cycles of equal order; and the ideal $\Phi(G; E)$ of the finite E-automorphisms of G is by Lemma 6.2 equal to $\mathfrak{A}$ so that in particular the identity is a finite automorphism. But this is the case if, and only if, G is the direct sum of a finite number of cycles, proving that (1) is a consequence of (2) as well as of (4).

If condition (1) is satisfied by the group G, then the maximum and the minimum condition are both satisfied by the E-admissible subgroups of G. Hence it follows from the Fundamental Theorem of section **2** that the maximum condition is satisfied by the right-annulets and by the left-annulets in $\mathfrak{A}(G; E)$. Furthermore it follows from Corollary 7.4 that annulets are principal ideals and that the sums of a finite number of principal ideals are annulets. Applying the maximum condition for the annulets it is readily deduced that every ideal is a principal ideal and therefore an annulet; and this shows that (3) and (5) are consequences of (1).

8. Products of Annulets

If $\mathfrak{T}$ and Ψ are right-(left-)ideals in the ring $\mathfrak{A}$, then their product $\mathfrak{T}\Psi$ is the smallest right-(left-)ideal in $\mathfrak{A}$ which contains all the products $v\psi$ for v in $\mathfrak{T}$ and ψ in Ψ. Annulets are ideals. Thus the product of two right-(left-)annulets is a well-determined right-(left-)ideal, though it need not be an annulet.

---

[37] For further characterizations of this class of groups, see Theorem 8.3 below and Asano (1), Satz 8, p. 243.

[38] Both finite and infinite sums.

[39] Rings with the properties (3) and (5) have been termed quasi-Frobeniusean by Nakayama (1).

THEOREM 8.1: *Suppose that G is a primary abelian operator group over the (primary) ring E, and that S and T are anti-cycles in $D(G; E)$.*
(a) *If either S is an anti-cycle of maximum index in $D(G; E)$ or T is a direct summand of G and $n(G/T) \leqq n(G/S)$, then* $\mathrm{P}(S)\mathrm{P}(T) = \mathrm{P}(S + GP^{n(G/T)})$
(b) *If both S and T are direct summands of G, but $n(G/S) < n(G/T)$, then* $\mathrm{P}(S)\mathrm{P}(T)$ *is not a right-annulet in* $\mathrm{A}(G; E)$.

PROOF: If σ and τ are E-automorphisms in $\mathrm{P}(S)$ and $\mathrm{P}(T)$ respectively, then $S \leqq N(\sigma) \leqq N(\sigma\tau)$, $G^{\sigma\tau} \leqq G^{\tau}$ and G^{τ} is a cycle in $D(G; E)$ whose order is at most $n(G/T)$. Hence $n(G^{\sigma\tau}) \leqq n(G^{\tau}) \leqq n(G/T)$ so that $GP^{n(G/T)}$ is part of $N(\sigma\tau)$, since $G/N(\sigma\tau)$ is a cycle of order $n(G^{\sigma\tau})$. The product $\sigma\tau$ is therefore an element in $\mathrm{P}(S + GP^{n(G/T)})$, i.e.

$$\mathrm{P}(S)\mathrm{P}(T) \leqq \mathrm{P}(S + GP^{n(G/T)}).$$

Since S and T are anti-cycles in $D(G; E)$, there exist elements v, w such that $G = S + vE = T + wE$. Clearly $n(G/S) \leqq n(v)$, $n(G/T) \leqq n(w)$. If one of the hypotheses of (a) is satisfied by S and T, then it is possible to select the element w in such a way that $n(w) \leqq n(G/S)$, since G is the sum of T and of a cycle of order $n(G/T)$, if T is a direct summand of G and since the maximum index of the anti-cycles in $D(G; E)$ is just the maximum order of the cycles in $D(G; E)$. Consequently there exists an E-automorphism σ in $\mathrm{P}(S)$ which maps v upon w. If φ is an E-automorphism in $\mathrm{P}(S + GP^{n(G/T)})$, then $S^{\varphi} = 0$ and $n(v^{\varphi}) \leqq n(G/T)$; and we note that φ is completely determined among the automorphisms in $\mathrm{P}(S)$ by the image of v. There exists one and only one E-automorphism τ in $\mathrm{P}(T)$ which maps w upon v^{φ}; and it is readily verified that $\varphi = \sigma\tau$, proving (a).

If both S and T are direct summands of G, then we may assume that the elements v and w selected above are of order $n(G/S)$ and $n(G/T)$ respectively. Denote by ϵ the E-automorphism in $\mathrm{P}(S)$ which leaves invariant the element v. If $n(G/S) < n(G/T)$, and if σ and τ are automorphisms in $\mathrm{P}(S)$ and $\mathrm{P}(T)$ respectively, then G^{σ} is a cycle in $D(G; E)$ whose order $n(G^{\sigma})$ is at most $n(G/S) < n(G/T)$ so that $T + G^{\sigma} \leqq T + GP$. Consequently $G^{\sigma\tau} \leqq (T + GP)^{\tau} \leqq GP$ though $G^{\epsilon} \not\leqq GP$. Since $\mathrm{P}(S)$ is the smallest right-annulet containing $\mathrm{P}(S)\mathrm{P}(T)$, we have thus shown that $\mathrm{P}(S)\mathrm{P}(T)$ is not a right-annulet.

THEOREM 8.2: *Suppose that G is a primary abelian operator group over the (primary) ring E, and that S and T are cycles in $D(G; E)$.*
(a) *If either T is a cycle of maximum order in $D(G; E)$ or S is a direct summand of G and $n(S) \leqq n(T)$, then* $\Lambda(S)\Lambda(T) = \Lambda(TP^{n(T)-n(S)})$.
(b) *If both S and T are direct summands of G, but $n(T) < n(S)$, then $\Lambda(S)\Lambda(T)$ is not a left-annulet in $\Lambda(G; E)$.*

PROOF: If σ and τ are E-automorphisms in $\Lambda(S)$ and $\Lambda(T)$ respectively, then G^{σ} is a subcycle of S and $G^{\sigma\tau}$ is therefore a subcycle of T whose order does not exceed $n(S)$ so that $G^{\sigma\tau} \leqq TP^{n(T)-n(S)}$ or

$$\Lambda(S)\Lambda(T) \leqq \Lambda(TP^{n(T)-n(S)}).$$

We may assume that $S \neq 0$. There exists a greatest E-admissible subgroup

Z of G whose cross-cut with S is 0. If Z^* is an E-admissible subgroup such that $Z < Z^*$, then Z^* contains the uniquely determined subcycle $SP^{n(S)-1}$ of S; and hence it follows from **1F** that Z is an anti-cycle in $D(G; E)$. If V is an anti-cycle in $D(G; E)$ whose cross-cut with S is 0 and whose index is as small as possible, and if one of the hypotheses of (a) is satisfied by S and T, then $n(G/V) \leqq n(T)$. There exists an E-automorphism τ of G such that $V^\tau = 0$ and G^τ is the subcycle $TP^{n(T)-n(G/V)}$ of order $n(G/V)$ of T; and τ clearly belongs to $\Lambda(T)$. Suppose now that φ belongs to $\Lambda(TP^{n(T)-n(S)})$. Then G^φ is a subcycle of T whose order does not exceed $n(S)$, and $G/N(\varphi)$ is therefore a cycle of an order not exceeding $n(S)$ and is generated by some coset C. Since a coset of G/V whose order does not exceed $n(S)$ contains one and only one element in S, there exists one and only one element v in S which is mapped by τ upon C^φ. There exists furthermore an E-automorphism σ in $\Lambda(S)$ which maps $N(\varphi)$ upon 0 and C upon v; and it is readily verified that $\varphi = \sigma\tau$, proving (a).

If $n(T) < n(S)$, and if S is a direct summand of G, then it is readily verified that $\Lambda(T)$ is the smallest left-annulet which contains $\Lambda(S)\Lambda(T)$. If T is a direct summand of G, then there exists an E-automorphism ω of G which maps G upon T and which leaves invariant every element in T. This automorphism ω is in $\Lambda(T)$, but it cannot be in $\Lambda(S)\Lambda(T)$. For if σ is in $\Lambda(S)$ and τ in $\Lambda(T)$, then T^σ is a subcycle of S whose order does not exceed $n(T) < n(S)$ so that $T^\sigma \leqq SP \leqq GP$, and hence $T^{\sigma\tau} \leqq TP$, though $T^\omega = T$; and this completes the proof.

THEOREM 8.3: *If G is a primary abelian operator group over the (primary) ring E, then each of the following properties implies the others*:

(a) *G is the direct sum of a finite number of cycles of equal order in $D(G; E)$.*

(b) *The product of any two right-annulets in $\Lambda(G; E)$ is a right-annulet.*

(c) *The product of any two left-annulets in $\Lambda(G; E)$ is a left-annulet.*

PROOF: That (a) implies both (b) and (c), is an immediate consequence of Theorem 7.5. If (b) or (c) is satisfied by G, then we deduce from Theorems 8.1, (b) and 8.2, (b) that G is the direct sum of cycles of equal order in $D(G; E)$, since G is, by **1E**, a direct sum of cycles in $D(G; E)$. If Z is any cycle of maximum-order in $D(G; E)$, then Z is a direct summand of G and there exists to every element g in G an E-automorphism of G which maps Z upon the cycle gE. Hence it follows from Lemma 6.2 that $\Lambda(Z)\Lambda(G) = \Lambda(Z)\Lambda = \Phi$, the ideal of all the finite automorphisms. Thus (c) implies that **1** is a finite E-automorphism, showing that (a) is a consequence of (c). If V is an anti-cycle of maximum-index in $D(G; E)$, then V is a direct summand of G and every anti-cycle in $D(G; E)$ may be mapped upon V by an E-automorphism of G. Hence we infer from Lemma 6.2 that $\mathrm{P}(0)\mathrm{P}(V) = \Lambda\mathrm{P}(V) = \Phi$. Consequently **1** is a finite E-automorphism, if (b) is satisfied by G; and (a) is therefore an implication of (b).

9. Cross-cuts of Right- and Left-Annulets

Right- and left-annulets in Λ are ideals in Λ; and they are therefore subrings of Λ. The cross-cut[40] of a right-annulet and a left-annulet in Λ is an ideal in

[40] The cross-cut of the sets S' and S'' shall always be denoted by $S' \cap S''$.

both these subrings of A. It is the object of this section to investigate the ideal theory obtained this way.

THEOREM 9.1: *Suppose that G is a primary abelian operator group over the (primary) ring E, and that H is an anti-cycle of index $n \neq 0$ in $D(G; E)$.*

(a) *If H is a direct summand of G, then every left-ideal I in the ring $\mathrm{P}(H)$ is the cross-cut of $\mathrm{P}(H)$ and of a left-ideal in $\mathrm{A}(G; E)$.*

(b) *If I is the cross-cut of $\mathrm{P}(H)$ and of some left-ideal in $\mathrm{A}(G; E)$, then there exists one and only one E-admissible subgroup S of G such that the orders of the elements in S do not exceed n and such that $I = \mathrm{P}(H) \cap \mathrm{A}(S)$.*

(c) *If the orders of the elements in the E-admissible subgroups S_v of G do not exceed n, then $\sum_v (\mathrm{P}(H) \cap \mathrm{A}(S_v)) = \mathrm{P}(H) \cap \mathrm{A}(\sum_v S_v)$.*

PROOF: If H is a direct summand of G, then there exists an (idempotent) E-automorphism η of G which maps H upon 0 and every coset C of G/H upon an element in C. If I^* is the smallest left-ideal in A containing I, then every element in I^* is the sum of elements $\varphi\chi$ for φ in A and χ in I. If ρ is an element in $\mathrm{P}(H)$, then $\rho = \eta\rho$; and thus it follows that the sum of elements $\varphi\chi$, if in $\mathrm{P}(H)$, is equal to the sum of elements $(\eta\varphi)\chi$. Since $\eta\varphi$ is in $\mathrm{P}(H)$, and since I is a left-ideal in the ring $\mathrm{P}(H)$, it follows that every $(\eta\varphi)\chi$ belongs to I and that therefore[41] $\mathrm{I} = \mathrm{I}^* \cap \mathrm{P}(H)$.

If I is the cross-cut of $\mathrm{P}(H)$ and of some left-ideal in A, then I is in particular the cross-cut of $\mathrm{P}(H)$ and I^*, the smallest left-ideal in A containing I. Since every element in I induces a linear transformation of the cycle G/H of order n upon some subcycle of G, it follows that the orders of the elements in $S = G^{\mathrm{I}}$ do not exceed n. If C is some coset which generates the cycle G/H, then S consists of all the elements $C^{\chi}e$ for χ in I and e in E. It is a consequence of Lemma 5.1 that there exists an E-automorphism φ of G which maps H into part of H and the coset C into part of Ce. Since $C^{\chi\varphi} = C^{\chi}e$, and since $\varphi\chi$ is in I^* and maps H upon 0, it follows that $\varphi\chi$ is in I and that S consists exactly of all the elements C^{ι} for ι in I. Thus S is an E-admissible subgroup of G the orders of whose elements do not exceed n. If κ is an element in $\mathrm{P}(H) \cap \mathrm{A}(S)$, then there exists an element ζ in I such that $C^{\zeta} = C^{\kappa}$, since S consists of all the C^{ι} for ι in I. But this implies clearly $\zeta = \kappa$ so that $\mathrm{I} = \mathrm{P}(H) \cap \mathrm{A}(S)$. Suppose now that T is an E-admissible subgroup of G the orders of whose elements do not exceed n and which satisfies: $\mathrm{I} = \mathrm{P}(H) \cap \mathrm{A}(T)$. If t is an element in T, then $n(tE) = n(t) \leqq n = n(G/H)$; and it is a consequence of **1E** that there exists an E-automorphism γ of G which maps H upon 0 and C upon t (γ clearly belongs to I). Hence $T = G^{\mathrm{I}} = S$.

If the orders of the elements in the E-admissible subgroups S_v of G do not exceed n, then the orders of the elements in the E-admissible subgroup $S = \sum_v S_v$ do not exceed n (by **1A**). Clearly $\sum_v (\mathrm{P}(H) \cap \mathrm{A}(S_v)) \leqq \mathrm{P}(H) \cap \mathrm{A}(S)$. If φ is an element in $\mathrm{P}(H) \cap \mathrm{A}(S)$, then C^{φ} for C a generating coset of G/H is an element in S so that $C^{\varphi} = s_1 + \cdots + s_k$ where s_i is an ele-

[41] The hypothesis that H is an anti-cycle in $D(G;E)$ has not been used in this part of the proof.

ment in S_{v_i}. There exists an E-automorphism φ_i in $\mathrm{P}(H) \cap \Lambda(S_{vi})$ which maps C upon s_i. Clearly $\varphi = \varphi_1 + \cdots + \varphi_k$, proving the desired equality (c).

THEOREM 9.2: *Suppose that G is a primary abelian operator group over the (primary) ring E, that Z is a cycle of order $n \neq 0$ in $D(G; E)$, and that I is a right-ideal in the subring $\Lambda(Z)$ of $\Lambda(G; E)$.*

(a) *If Z is a direct summand of G, then I is the cross-cut of $\Lambda(Z)$ and of a right-ideal in Λ.*[42]

(b) *There exists at most one E-admissible subgroup S of G such that the orders of the elements in G/S do not exceed n and such that $\mathrm{I} = \Lambda(Z) \cap \mathrm{P}(S)$.*

(c) *If I is the cross-cut of $\Lambda(Z)$ and of a right-ideal in Λ, and if $G/N(\mathrm{I})$ is a sum of a finite number of cycles, then the orders of the elements in $G/N(\mathrm{I})$ do not exceed n and $\mathrm{I} = \Lambda(Z) \cap \mathrm{P}(N(\mathrm{I}))$.*

PROOF: If Z is a direct summand of G, then there exists an (idempotent) E-automorphism ζ of G which maps G upon Z and which leaves invariant all the elements in Z. The smallest right-ideal in Λ which contains I is $\mathrm{I}^* = \mathrm{I}\Lambda$, and every element in I^* is the sum of products $\chi\varphi$ for χ in I and φ in Λ. If ω is an element in $\Lambda(Z)$, then $\omega = \omega\zeta$; and every element in $\mathrm{I}^* \cap \Lambda(Z)$ is therefore the sum of products $\chi(\varphi\zeta)$ for χ in I and φ in Λ. Since every $\varphi\zeta$ is in $\Lambda(Z)$, and since I is a right-ideal in $\Lambda(Z)$, it follows that every $\chi\varphi\zeta$ is in I, i.e. $\mathrm{I} = \Lambda(Z) \cap \mathrm{I}^*$.

The statement (b) is an immediate consequence of the following lemma.

(b*) If S is an E-admissible subgroup of G, if the orders of the elements in G/S do not exceed n, then $S = N(\Lambda(Z) \cap \mathrm{P}(S))$.

It is clear that $S \leqq N(\Lambda(Z) \cap \mathrm{P}(S))$. To prove the opposite inequality consider a basis B of G/S. Since the orders of the elements in B do not exceed n, there exists to every element b in B an E-automorphism $\varphi(b)$ of G in $\mathrm{P}(S)$ which maps the coset b of G/S upon an element of order $n(b)$ in Z and which maps all the other elements in B upon 0. If x is an element not 0 in G/S, then $x = \sum_{b \text{ in } B} be(b)$ where the $e(b)$ are elements in E almost all of which are 0. There exists at least one element in B, say v, such that $ve(v) \neq 0$; and it is clear that $x^{\varphi(v)}$ is different from 0. This completes the proof of (b*) (and (b)).

If γ is any automorphism in $\Lambda(Z)$, then $G/N(\gamma)$ is a cycle of an order not exceeding n and consequently $GP^n \leqq N(\gamma)$. If I is any subset of $\Lambda(Z)$, then $N(\mathrm{I})$ is the cross-cut of all the $N(\gamma)$ for γ in I so that $GP^n \leqq N(\mathrm{I})$; and the orders the of elements in $G/N(\mathrm{I})$ do not exceed n.

If I is the cross-cut of $\Lambda(Z)$ and of some right-ideal in Λ, then $\mathrm{I} = \Lambda(Z) \cap \mathrm{I}^*$ where $\mathrm{I}^* = \mathrm{I}\Lambda$ is the smallest right-ideal in Λ which contains I. It is clear that $N(\mathrm{I}) = N(\mathrm{I}^*)$ and that $\mathrm{I} \leqq \Lambda(Z) \cap \mathrm{P}(N(\mathrm{I}))$. We consider the set Q of all the pairs (U, V) with the following properties:

(i) U and V are E-admissible subgroups of $G/N(\mathrm{I})$ and $G/N(\mathrm{I})$ is the direct sum of U and V.

(ii) To every automorphism $\mathfrak{s}$ in $\Lambda(Z) \cap \mathrm{P}(N(\mathrm{I}))$ there exists an automorphism $\mathfrak{s}^*$ in I which maps V upon 0 and which coincides with $\mathfrak{s}$ on U.

---

[42] For the proof of (a) we need not assume that Z is a cycle in $D(G;E)$.

This set Q is not vacuous, since the pair $(0, G/N(\mathrm{I}))$ meets the requirements (i), (ii). To prove (c) we assume that $G/N(\mathrm{I})$ is the sum of a finite number of cycles. Thus the maximum condition is satisfied by the E-admissible subgroups of $G/N(\mathrm{I})$ and hence there exists a pair (U^*, V^*) in Q whose first member U^* is as big as possible To prove (c) we need only show that $V^* = 0$. If V^* were not 0, then V^* would contain a subcycle W of maximum order in V^*. To every element, not 0, in W there exists an automorphism in I which maps this element upon an element, not 0, in Z (as follows from the definition of $N(\mathrm{I})$ and from the fact that $\mathrm{I} \leqq \Lambda(Z)$). Thus there exists an automorphism δ in I which maps the cycle $WP^{n(W)-1}$ of order 1 upon the cycle ZP^{n-1} of order 1; and it is readily seen that $W^\delta = ZP^{n-n(W)}$ is a cycle of order $n(W)$. Since (U^*, V^*) belongs to Q, there exists an automorphism δ^* in I which coincides with δ in U^* and which maps V^* upon 0. The automorphism $\epsilon = \delta - \delta^*$ belongs to I (since I is an ideal), maps U^* upon 0 and W upon $ZP^{n-n(W)}$. Denote by V^{**} the set of elements in V^* which are mapped upon 0 by ϵ, and put $U^{**} = U^* + W$. Since ϵ maps V^* upon a cycle of order $n(W)$, and since W is a subcycle of maximum order of V^*, it follows that $G/N(\mathrm{I})$ is the direct sum of the E-admissible subgroups U^{**} and V^{**}. If φ is any automorphism in $\Lambda(Z) \cap \mathrm{P}(N(\mathrm{I}))$, then there exists—by (ii)—an automorphism φ^* in I which coincides with φ on U^* and which maps V^* upon 0. Since φ maps W upon part of $ZP^{n-n(W)} = W^\epsilon$, there exists an E-automorphism ρ of $ZP^{n-n(W)}$ such that φ and $\epsilon\rho$ coincide on W. There exists by Lemma 5.1 an E-automorphism σ of G which induces ρ in $ZP^{n-n(W)}$; and $\epsilon\sigma$ is clearly an automorphism in $\mathrm{I} = \mathrm{I}^* \cap \Lambda(Z)$. The automorphism $\varphi^{**} = \varphi^* + \epsilon\sigma$ belongs to I, maps V^{**} upon 0 and coincides with φ on U^{**}. Thus we have shown that the pair (U^{**}, V^{**}) belongs to the set Q in spite of the fact that $U^* < U^{**}$, a contradiction which completes the proof.

COROLLARY 9.3: *Suppose that G is a primary abelian operator group over the (primary) ring E, and that Z is a cycle of order $n(\neq 0)$ in $D(G; E)$.*
(a) *G/GP^n is the sum of a finite number of cycles if, and only if, $\Lambda(Z) \cap \mathrm{I} = \Lambda(Z) \cap \mathrm{P}(N(\Lambda(Z) \cap \mathrm{I}))$ for every right-ideal I in $\Lambda(G; E)$.*
(b) *If S and T are E-admissible subgroups of G such that the orders of the elements in G/S and in G/T do not exceed n, then $(\Lambda(Z) \cap \mathrm{P}(S)) + (\Lambda(Z) \cap \mathrm{P}(T)) = \Lambda(Z) \cap \mathrm{P}(S \cap T)$.*

PROOF: $GP^n \leqq N(\Lambda(Z) \cap \mathrm{I})$, since Z is a cycle of order n. If G/GP^n is the sum of a finite number of cycles, then $G/N(\Lambda(Z) \cap \mathrm{I})$ is the sum of a finite number of cycles; and $\Lambda(Z) \cap \mathrm{I} = \Lambda(Z) \cap \mathrm{P}(N(\Lambda(Z) \cap \mathrm{I}))$ for right-ideals I in Λ is a consequence of Theorem 9.2, (c). Suppose now that this identity holds for every right-ideal I in Λ. It is a consequence of **1E** that G/GP^n is the direct sum of cycles Z_v of positive order not exceeding n. The set I of all the E-automorphisms which map GP^n and almost all the Z_v upon 0 is readily seen to be a right-ideal in Λ; and the cross-cut $\mathrm{I} \cap \Lambda(Z)$ consists exactly of those E-automorphisms of G which map G into part of Z and which map GP^n and almost all the Z_v upon 0. Since $n(Z_v) \leqq n = n(Z)$, it follows that $N(\mathrm{I}) = GP^n =$

$N(\Lambda(Z) \cap \mathfrak{I})$. Hence there exists an automorphism ω in $\mathfrak{P}(N(\Lambda(Z) \cap \mathfrak{I}))$ which maps GP^n upon 0 and [every] Z_v upon $ZP^{n-n(Z_v)}$. Since ω belongs to $\Lambda(Z)$, it follows from the validity of the identity in (a), that ω belongs to $\mathfrak{I}$ so that the automorphism ω which maps none of the Z_v upon 0 maps almost all the Z_v upon 0. Hence there exists only a finite number of Z_v ; and this completes the proof of (a).

To prove (b) we need only show that every automorphism φ in $\Lambda(Z) \cap \mathfrak{P}(S \cap T)$ belongs to $(\Lambda(Z) \cap \mathfrak{P}(S)) + (\Lambda(Z) \cap \mathfrak{P}(T))$. Since the orders of the elements in G/S and G/T do not exceed n, it follows that GP^n is part of S and of T, and that therefore the orders of the elements in $G/(S \cap T)$ do not exceed n. Hence an E-automorphism γ of G may be constructed which maps S upon 0 and coincides with φ on T and which maps G into part of Z. Then τ is in $\Lambda(Z) \cap \mathfrak{P}(S)$ and $\varphi - \tau$ is in $\Lambda(Z) \cap \mathfrak{P}(T)$; and this completes the proof of (b).

REMARK: In sections **5** to **8** we obtained fairly similar results for right- and for left-annulets, though we had to use different methods in proving these facts. In this section **9** even the results turned out to be different (cf. in particular Corollary 9.3, (a) and Theorem 9.1, (c)). The following argument will give some explanation of this difference in behaviour between right- and left-annulets.

1. If Z is a cycle of order $n \neq 0$ in $D(G; E)$, E a primary ring, then GP^n is mapped upon 0 by all the automorphisms in $\Lambda(Z)$; and the automorphisms in $\Lambda(Z)$ map G/GP^n into the cycle Z. Thus the automorphisms in $\Lambda(Z)$ behave just like the characters of G/GP^n with values in Z; and $\Lambda(Z)$ is essentially the character group of G/GP^n (in Z).[43] If S is an E-admissible subgroup of G which contains GP^n, then $\Lambda(Z) \cap \mathfrak{P}(S)$ consists just of those characters of G/GP^n which map S upon 0. This correspondence between the subgroups in $D(G/GP^n; E)$ and right-ideals in the ring $\Lambda(Z)$ is a duality of $D(G; E)$ upon a certain system of right-ideals in the ring $\Lambda(Z)$.

2. If H is an anti-cycle of index $n \neq 0$ in $D(G; E)$, C a generating coset of the cycle G/H, and G_n the subgroup of all the elements in G whose order does not exceed n, then there exists to every element x in G_n one and only one E-automorphism $\varphi = \varphi(x)$ in $\mathfrak{P}(H)$ such that $x = C^{\varphi}$, and hence $C^{\mathfrak{P}(H)} = G_n$. Mapping x upon $\varphi(x)$ constitutes an isomorphism of G_n upon $\mathfrak{P}(H)$; and $\mathfrak{P}(H) \cap \Lambda(S)$ is just the image of the E-admissible subgroup S of G under this isomorphism. Thus G_n and $\mathfrak{P}(H)$ are essentially the same.

3. If the group G is the sum of a finite number of cycles, then it is well known[44] that G and its character group are very similar and their respective systems of subgroups are duals. Without this hypothesis this duality breaks down[45] and can only be reconstructed by resort to topological means.[46] Since these do not

---

[43] Cp. Lewis (1).

[44] Cp. e.g. Baer (4), II, 4 and Lewis (1).

[45] It has been shown that this finiteness-hypothesis is (under certain circumstances) actually necessary for the existence of a dual; cp. Baer (3).

[46] Pontrjagin (1), Ch. V.

enter into our considerations, G and its character group will in general be rather dissimilar; and thus right- and left-annulets will behave differently.

Chapter IV: Characterization of Automorphism Rings

In Chapter III we have derived a number of properties which are satified by the annulets in the automorphism rings of primary abelian operator groups. It is the object of this present Chapter IV to prove that the automorphism rings of primary abelian operator groups are the only rings with these properties.

10. Idempotents

We consider a ring K which shall always be assumed to contain an identity element 1. If S is any subset of K, then we define as the right (resp. left)-annulet $R(S)$ (resp. $L(S)$) determined by S the set of all the elements x in K which satisfy: $sx = 0$ (resp. $xs = 0$) for every s in S. The right (resp. left)-annulets are right (resp. left)-ideals; cross-cuts of right (resp. left)-annulets are right (resp. left)-annulets; $S \leqq T$ implies $R(T) \leqq R(S)$ (resp. $L(T) \leqq L(S)$); and $R(L(R(S))) = R(S)$ (resp. $L(R(L(S))) = L(S)$).

If U and V are subsets of K, then their sum $U + V$ consists of all the sums $u + v$ for u in U and v in V. If U and V are ideals whose cross-cut is 0, then $U + V$ is the direct sum of U and V, and U is a direct summand of $U + V$. An element g in K is an idempotent if $g = g^2$. In the following we collect certain well known[47] connections between the concepts: annulet, direct summand of the ring K and idempotent, which will prove useful in the future; but it should be kept in mind that these remarks refer only to an important subclass of the set of annulets.

Lemma 10.1: *The ring K is the direct sum of the right (resp. left)-ideals U and V if, and only if, there exist idempotents u and v such that $1 = u + v$, $uv = vu = 0$ and $U = uK$, $V = vK$ (resp. $U = Ku$, $V = Kv$).*

Lemma 10.2: *If the idempotents u and v in the ring K satisfy: $1 = u + v$ and $uv = vu = 0$, then $R(u) = R(Ku) = vK$, $R(v) = R(Kv) = uK$ (resp. $L(u) = L(uK) = Kv$, $L(v) = L(vK) = Ku$).*

Lemma 10.3: *The right (resp. left)-ideal U is a direct summand of K if, and only if, there exists an idempotent u such that $U = uK$ (resp. $U = Ku$).*

Lemma 10.4: *If u is an idempotent and S a subset of the ring K which contains uS and/or Su, then $uS = S \cap uK$ and/or $Su = S \cap Ku$.*

We note in particular the following important consequence of Lemmas 10.1 and 10.2: If K is the direct sum of the right (resp. left)-ideals U and V, then U and V are right (resp. left)-annulets whose cross-cut is 0; and K is the direct sum of $L(U)$ and $L(V)$ (resp. $R(U)$ and $R(V)$) so that $L(U) \cap L(V) = 0$ (resp. $R(U) \cap R(V) = 0$).

11. Groups Contained in the Ring K

If J is a right-ideal in the ring K, then J is a subring of K and in particular J

[47] Cp. e.g. von Neumann (1), Ch. II, p. 7/8.

is an additive abelian group. Every cross-cut of J and of a left-ideal in K is a subgroup of this abelian group J and *the set of all these subgroups* shall be denoted by $D(J)$.

If x is any element in K satisfying $xJ \leqq J$, then $xS \leqq S$ for every subgroup S of the group J which is contained in $D(J)$; the elment x induces in J an endomorphism (cf. section **4**) which may be termed $D(J)$-*admissible*, since it maps every subgroup in $D(J)$ into itself. *The system of all the endomorphisms of J which are obtained by left-multiplication by elements in K* shall be denoted by $E(J)$. The elements x in K which satisfy: $xJ \leqq J$ form a subring $L^*(J)$ of K which clearly contains $L(J)$; and one verifies readily that mapping the element x in $L^*(J)$ upon the endomorphism in $E(J)$ which x induces in J constitutes an anti-isomorphism of $L^*(J)/L(J)$ upon $E(J)$.

If x is any element in K, then $Jx \leqq J$, since J is a right-ideal in K and the element x induces therefore in this fashion an endomorphism in the group J. If e is any endomorphism in $E(J)$, then e maps every element y in J upon an element ye in J; and there exists by the definition of $E(J)$ an element f in $L^*(J)$ such that $fy = ye$ for every y in J. If x is any element in K, y an element in J, e an element in $E(J)$ and f an element in $L^*(J)$ inducing e in J, then $(ye)x = (fy)x = f(yx) = (yx)e$ so that x *induces in J an $E(J)$-automorphism* which we denote by y^x. *The system of these $E(J)$-automorphisms of J* shall be denoted by $A(J)$; and it is clearly isomorphic to $K/R(J)$ (this is not important, since in all our applications $R(J)$ will be 0).

LEMMA 11.1: *If the right-ideal J in K is generated by some idempotent j, i.e. if $J = jK$, then every endomorphism in $E(J)$ is induced by one and only one element in the ring jKj (with identity element j) and in this fashion an anti-isomorphism of the ring jKj upon the ring $E(J)$ is defined; and $D(J)$ is exactly the system $D(J; E(J))$ of all the $E(J)$-admissible subgroups of J as well as the system of all the left-ideals in the ring J.*

PROOF: An element y belongs to the right-ideal $J = jK$, for j an idempotent, if, and only if, $y = jy$. If e is any element in $E(J)$, then there exists an element f in $L^*(J)$ such that $ye = fy$ for every y in J. Since y and ye are elements in J, we have $y = jy$ and $ye = j(ye)$; and thus it follows that $ye = j(ye) = j(fy) = (jf)y = (jf)(jy) = (jfj)y$. It is obvious that $jKj \leqq L^*(J)$ and that jfj induces the null-element of $E(J)$ if, and only if, $jfj = 0$.

The cross-cut of J and of a left-ideal in K is a left-ideal in the ring J. From what we have shown just now it follows that every left-ideal in the ring J is an $E(J)$-admissible subgroup of the abelian group J. If finally S is an $E(J)$-admissible subgroup of the abelian group J, then $S = jS$, since $S \leqq jK = J$, and $S = SE(J)$. Consequently we may deduce from Lemma 10.4 that the cross-cut of the left-ideal KS and of J is exactly $KS \cap J = jKS = jKjS = SE(J) = S$; and this completes the proof of our lemma.

12. The Postulates

In this section we are going to enumerate a number of properties of rings; we shall show that the automorphism rings of primary abelian operator groups

meet these requirements and we shall discuss a few of the interrelations between these postulates. In the next section we shall prove that the automorphism rings of primary abelian operator groups are the only rings satisfying these conditions.

If K is a ring (containing an identity element 1), then we shall denote by $RA(K)$ (resp. $LA(K)$) *the partially ordered set of all the right (resp. left)-annulets in* K. Both $RA(K)$ and $LA(K)$ contain the (set-theoretical) cross-cuts of all their subsets; and thus they contain with any subset a smallest annulet containing all the annulets in the given set: the *join* of the annulets in the set. We note that these joins of annulets need not be their ideal-theoretical sums. The mappings: H upon $L(H)$ and J upon $R(J)$ constitute reciprocal dualities of $RA(K)$ upon $LA(K)$ and of $LA(K)$ upon $RA(K)$.

I. *If* $H_i = R(J_i)$ *and* $J_i = L(H_i)$, *and if* $H_1 \cap H_2 = 0 = J_1 \cap J_2$, *then* $H_1 + H_2 = K = J_1 + J_2$.

It is a consequence of Theorem 7.1, (c) that this postulate I is satisfied by the automorphism rings Λ of primary abelian operator groups.

II. *If* H *is a cycle of order* $n \neq 0$ *in* $RA(K)$, *then*

(a) $D(H)$ *is a ring*[48] *of subgroups of the additive abelian group* H; *and*

(b) *to every* S *in* $D(H)$ *there exists one and only one* S^* *in* $LA(K)$ *such that* $S = H \cap S^*$ *and such that the orders of the subcycles of* S^* *in* $LA(K)$ *do not exceed* n.

It is a consequence of Theorem 9.1, (c), (b) that this postulate II is satisfied by the automorphism rings of primary abelian operator groups.

III. *If* H *is a cycle of maximum order* m *in* $RA(K)$, *if* J *is a cycle of order* $n(\leqq m)$ *in* $LA(K)$, *if* K *is the direct sum of* J *and some left-annulet in* K, *and if* V *is a right-ideal in* K *such that* $V \cap J \leqq H$, *then there exists a right-annulet* W *in* K *such that* $W \cap J = V \cap J$.

That III is satisfied by the automorphism rings $\Lambda(G; E)$ of primary abelian operator groups G over (primary) rings E, may be seen as follows: It is a consequence of the Fundamental Theorem of section **2** that H is a cycle of maximum order m in $RA(\Lambda)$ if, and only if, $\mathrm{H} = \mathrm{P}(S)$ for S an anti-cycle of maximum index m in $D(G; E)$; and that I is a cyclic direct summand of Λ in $LA(\Lambda)$ if, and only if, $\mathrm{I} = \Lambda(T)$ for T a cyclic direct summand of G in $D(G; E)$. If Ψ is a right-ideal in Λ such that $\Psi \cap \mathrm{I} \leqq \mathrm{H}$, then $N(\mathrm{H}) \leqq N(\Psi \cap \mathrm{I})$. Since $S = N(\mathrm{P}(S)) = N(\mathrm{H})$, it follows that $N(\Psi \cap \mathrm{I})$ is an anti-cycle in $D(G; E)$; and that therefore by Theorem 9.2, (c) we have $\Psi \cap \mathrm{I} = \Lambda(T) \cap \mathrm{P}(N(\Psi \cap \mathrm{I})) = \mathrm{P}(N(\Psi \cap \mathrm{I})) \cap \mathrm{I}$, as was to be shown.

IV, i. *There exists an integer* m *such that both* $RA(K)$ *and* $LA(K)$ *contain at least* i *independent cycles of order* m *and such that neither contains cycles of an order exceeding* m.

This condition IV, i shall only be used for $i = 1, 2, 3$. We note that condition IV, 1 states only the following fact: The orders of the cycles in $RA(K)$ and in $LA(K)$ are bounded and the maximum order of the cycles in $RA(K)$ is the same as the maximum order of the cycles in $LA(K)$. It is a consequence of **1E** and

[48] In the sense of **1B**.

of the Fundamental Theorem of section **2** that postulate IV, 1 is satisfied by the automorphism rings of primary abelian operator groups. It should be noticed that III would be vacuous, if IV, 1 were not true.

THEOREM 12.1: (1) *If* II, (b) *and* IV, 1 *are satisfied by the ring* K, *and if* G *is a cycle of maximum order* m *in* $RA(K)$, *then a projection of* $LA(K)$ *upon* $D(G)$ *is effected by mapping the left-annulet* J *upon* $J \cap G$.
(2) *If* II *and* IV, 1 *are satisfied by the ring* K, *then* $RA(K)$ *and* $LA(K)$ *are complete modular lattices.*

PROOF: Every element in $D(G)$ has the form (by definition of $D(G)$): $G \cap V$ for V a left-ideal in K. Since m is the maximum order of the cycles in $LA(K)$, there exists to every left-ideal V in K one and only one left-annulet V^* such that $G \cap V = G \cap V^*$; and this proves (1). (2) is an immediate consequence of (1), since rings of subgroups of abelian groups are complete modular lattices, and since $RA(K)$ and $LA(K)$ are duals of each other

V. *The following conditions are satisfied by the partially ordered set* $LA(K)$.
(a) *The modular (Dedekind's) law.*
(b) *If* S *is a not vacuous set of cycles in* $LA(K)$, *if* S *contains every cycle in* $LA(K)$ *which is part of the join of a finite number of cycles in* S, *then there exists one and only one left-annulet* L_S *in* K *such that the cycles in* S *and only these are part of* L_S.
(c) *If* U *and* V *are left-annulets in* K, *if* $U \leqq V$, *and if there exist at most two different left-annulets between* U *and* V *which are modulo*[49] U *cycles of order* 1, *then* V *is a cycle modulo* U.

It is a consequence of **1E**, **1F** and the Fundamental Theorem of section **2** that this postulate is satisfied by the automorphism rings of primary abelian operator groups.

It has already been shown (Theorem 12.1, (2)) that V, (a) may be deduced from II and IV, 1. It is easily deduced from V, (b) that *every left-annulet in* K *is the join of cycles*[50] *in* $LA(K)$; and it follows from Theorem 12.1, (1) that V, (b) is a consequence of II, IV, 1 and this last condition.

The part of IV, 1 which states that the maximum order of the cycles is the same in $RA(K)$ and in $LA(K)$ may be derived from V, since $LA(K)$ and $RA(K)$ are duals of each other, and since it may be shown[51] that K is in $RA(K)$ and in $LA(K)$ the direct sum of a cycle of maximum order and of an anti-cycle of maximum index.

THEOREM 12.2: *If* I, II, (b), IV, 1 *and* V *are satisfied by the ring* K, *and if* G *is a cycle of maximum order* m *in* $RA(K)$, *then* $G = jK$ *for some idempotent* j *in* K, *and* $D(G)$ *is a ring of subgroups of* G.

PROOF: It is known[51] that there exists because of V an element H in $RA(K)$ satisfying: K is the join and 0 the cross-cut of G and H. This implies that 0 is the cross-cut of $L(G)$ and $L(H)$; and hence it follows from I that K is the

[49] U is modulo V a cycle of order n, if V is part of U, and if U is a cycle of order n in the partially ordered set of the elements between U and V.

[50] Though it need not be the join of a finite number of cycles.

[51] Baer (4), Theorem I.3.6 and I.5.1.

direct sum of G and H. Now we deduce from Lemma 10.1 that $G = jK$ for j an idempotent in K; and we deduce from Lemma 11.1 that $D(G)$ is just the ring of all the $E(G)$-admissible subgroups of the abelian group G.

VI. *If X and Y are cycles in $RA(K)$, and if X is a cycle of maximum order in $RA(K)$, then*[52] $XY = X^{(n(X)-n(Y))}$.

It is a consequence of Theorem 8.1, (a) that VI is satisfied by the automorphism rings of primary abelian operator groups.

The last postulate which we require will be vacuous, if K is the join of a finite number of cycles in $RA(K)$. To enunciate postulate VII two concepts are needed: We denote by $W = W(K)$ *the class of all the right-annulets in K which are cross-cuts of a finite number of anti-cycles in $RA(K)$*; and we term *W-function* any single-valued function $f(J)$ of the right-annulets J in the class W with the following properties:

$$f(J) \text{ is a coset of } K/J; \quad \text{and } J \leqq H \text{ implies } f(J) \leqq f(H).$$

We note that 0 is in the class W if, and only if, 0 is the cross-cut of a finite number of anti-cycles in $RA(K)$; and this is equivalent to saying that K is the join of a finite number of cycles in $RA(K)$.

VII. *To every W-function $f(J)$ there exists one and only one element f in K such that $f(J) = J + f$ for every J in W.*

It is a consequence of Theorem 6.3 that the postulate VII is satisfied by the automorphism rings of primary abelian operator groups.

It may be mentioned that the uniqueness of the element f in K occurring in VII may be derived from the other postulates.

13. Completeness of the System of Postulates

In this section we are going to prove the following *Existence Theorem*:

The postulates I to III, IV, 2, and V to VII are satisfied by the ring K containing an identity element 1 if, and only if, there exists a primary abelian operator group G over a (primary) ring E such that $D(G;E)$ contains at least two independent cycles of maximum order and such that K is isomorphic to the ring $A(G;E)$ of all the E-automorphisms of G.

It has been shown in section **12** that the postulates I to III, IV, 1, V to VII are satisfied by the automorphism rings of primary abelian operator groups; and that IV, 2 is satisfied whenever $D(G;E)$ contains at least two independent cycles of maximum order, is an immediate consequence of the Fundamental Theorem of section **2**.

If IV, 1 is satisfied by the ring K, then there exists a cycle G of maximum order m in the set $RA(K)$ of all the right-annulets in K. We are going to prove that

G is a primary abelian operator group over the (primary) ring $E(G)$, that $D(G;E(G))$ contains at least two independent cycles of the maximum order m and

[52] For these notations cp. **1A** and section **8.**

that K is (essentially) the ring of all the $E(G)$-automorphisms of the abelian group G, provided I, II, b, III, IV, 2, V to VII are satisfied by the ring K.

The proof of this theorem will be effected in several steps; and we shall make a note at each of these lemmas which of the hypotheses we actually applied in its proof.

13.1. *[$G = jK$ for some idempotent j in K;] $D(G)$ is the ring of all the left-ideals in the ring G and $D(G) = D(G;E(G))$; every endomorphism in $E(G)$ is induced by left-multiplication of G by one and only one element in the ring jKj; the rings jKj and $G/(L(G) \cap G)$ are isomorphic and the rings $E(G)$ and jKj are anti-isomorphic.* (I; II, b; IV, 1; V.)

This statement is an immediate consequence of Theorem 12.2 and of Lemma 11.1.

G is a cycle of order m in $RA(K)$ and hence it contains one and only one subcycle $G^{(i)}$ of order $m - i$ in $RA(K)$ for $0 \leqq i \leqq m$.

13.2. *$G^{(i)}$ is a two-sided ideal in the subring G of K* (VI).

PROOF: $G^{(i)}$ is a right-annulet and therefore a right-ideal in K. Since G is a cycle of maximum order m in $RA(K)$, it follows from VI that $GG^{(i)} = G^{(i)}$ so that $G^{(i)}$ is a left-ideal in the ring G.

13.3. *$G^{(m-i)}$ is the sum of all the cycles of an order not exceeding i in $D(G)$* (I; II, b; IV, 1; V; VI).

PROOF: Since $G^{(m-i)}$ is a cycle of order i in $RA(K)$, and since $G^{(m-i)}$ belongs to $D(G^{(m-i)})$, there exists one and only one left-annulet L in K which contains $G^{(m-i)}$ and which contains only cycles in $LA(K)$ of an order not exceeding i. If C is any cycle in $LA(K)$ whose order does not exceed i, then the orders of the subcycles of the join of L and C in $LA(K)$ do not exceed i (by **1A**). Hence C is part of L, since $G^{(m-i)}$ is part of the join of L and C. Clearly L is the smallest left-annulet containing $G^{(m-i)}$; and it is a consequence of V, b that L is the join of all the cycles in $LA(K)$ whose orders do not exceed i.

It is a consequence of 13.2 and 13.1 that $G^{(m-i)}$ belongs to $D(G)$. Hence there exists by II, b one and only one left-annulet L^* such that $G^{(m-i)} = G \cap L^*$. Since L is the smallest left-annulet containing $G^{(m-i)}$, this implies $L = L^*$. It is a consequence of II, b that a projectivity of $LA(K)$ upon $D(G)$ is obtained by mapping every left-annulet in K upon its cross-cut with G. Hence we may deduce from 13.1 that $G^{(m-i)} = G \cap L$ is the sum of all the cycles in $D(G)$ of an order not exceeding i.

13.4. *$E(G)$ is a cycle of order m in the partially ordered set of all the right-ideals*[53] *in the ring $E(G)$* (I; II, b; IV, 1; V).

PROOF: It is a consequence of 13.1 that the rings $G/(L(G) \cap G)$ and $E(G)$ are anti-isomorphic; and hence it suffices to prove that G is a cycle of order m in the partially ordered set of left-ideals in G which contain $L(G) \cap G$. If J is such a left-ideal in G, then it follows from 13.1 that J belongs to $D(G)$; and thus we may deduce from II, c that there exists one and only one left-annulet L

[53] This implies that every right-ideal in $E(G)$ is a two-sided ideal.

satisfying $J = G \cap L$. It is a consequence of II, b and of $L(G) \cap G \leqq L \cap G$, that $L(G) \leqq L$; and this implies $R(L) \leqq R(L(G)) = G$, since G is a right-annulet. But the $G^{(i)}$ for $0 \leqq i \leqq m$ are the only right-annulets contained in G so that $R(L) = G^{(i)}$, $L = L(G^{(i)})$ and $J = L(G^{(i)}) \cap G$ for some i. G is therefore a cycle of an order not exceeding m in the partially ordered set of left-ideals in G which contain $L(G) \cap G$. The order is exactly m, since $G^{(i)} \neq G^{(j)}$ for $0 \leqq i < j \leqq m$, since therefore all the $L(G^{(i)})$ are different, and since consequently the inequality of the $G \cap L(G^{(i)})$ for $0 \leqq i \leqq m$ may be deduced from II, b.

13.5. *$E(G)$ is a primary ring (with $m(E(G)) = m$)* (I; II, b; IV, 2; V).

PROOF: $E(G)$ contains an identity element 1 and G is an abelian operator group over the ring $E(G)$. It is a consequence of 13.1 that $D(G) = D(G;E(G))$; and it follows from IV, 2 that $D(G;E(G))$ contains at least two independent cycles of maximum order m. We infer from 13.4 and V, c that the criterion **1F** may be applied on the ring $E(G)$ and $E(G)$ is shown to be a primary ring with $m(E(G)) = m$.

We recall that $A(G)$ is the set of all the $E(G)$-automorphisms of the abelian group G which are induced by right-multiplication of G by elements in K; and that $A(G)$ and $K/R(G)$ are isomorphic rings.

13.6. *$R(G) = 0$ so that $A(G)$ and K are essentially the same* (II, b; IV, 1).

PROOF: There exists by II, b one and only one left-annulet containing G. But G is part of both K and $L(R(G))$ so that $K = L(R(G))$ or $0 = R(K) = R(G)$.

13.7. *If the abelian group G is the direct sum of the subgroup U in $D(G)$ and of the cycle $Z = zE(G)$ in $D(G)$, and if the order of the element t in G does not exceed $n(Z)$, then there exists one (and only one) $E(G)$-automorphism in $A(G)$ which maps U upon 0 and z upon t* (I; II, b; III; IV, 2; V; VI).

We note that it is a consequence of 13.1 that the cycles in $D(G)$ are of the form $gE(G)$ for g in G, and that it is a consequence of 13.5 that G (as well as the subgroups in $D(G)$) are primary abelian operator groups over the (primary) ring $E(G)$.

PROOF: There exists by II, b to every S in $D(G)$ one and only one left-annulet S^* in K such that $S = G \cap S^*$, since G is a cycle of maximum order m in $RA(K)$, and since m is by IV, 1 the maximum order of the cycles in $LA(K)$. It is a consequence of 13.1 that a projectivity of $D(G)$ upon $LA(K)$ is effected by mapping S upon S^*.

U is a primary abelian operator group over the (primary) ring $E(G)$. Hence it follows from **1E** that U is the direct sum of cycles Z_v in $D(G)$; and there exist uniquely determined elements c, c_v in Z and Z_v respectively almost all of which are 0, whose sum is t and whose orders do not exceed $n(t) \leqq n(Z)$.

K is the direct sum of Z^* and U^*, since 0 is the cross-cut and K the join of Z^* and U^* in $LA(K)$, since therefore 0 is the cross-cut of $R(Z^*)$ and $R(U^*)$, and since we may apply I. We infer from Lemma 10.1 the existence of idempotents e and f such that $1 = e + f$, $ef = fe = 0$, $Z^* = Ke$, $U^* = Kf$ and $R(U^*) = eK$. Likewise we prove the existence of idempotents e_v such that $Z_v^* = Ke_v$.

Since G is a cycle of maximum order in $RA(K)$, since Z^* is a cycle in $LA(K)$, since K is the direct sum of Z^* and of some left-annulet in K, and since $zK \cap Z^* = zK \cap Ke = zKe \leqq G$ (by Lemma 10.4 and the right-ideal property of G), it follows from III that there exists a right-annulet T in K such that $zKe = T \cap Z^*$. Since the cross-cut of right-annulets is a right-annulet, and since G is a cycle in $RA(K)$, we deduce that $T \cap G = G^{(i)}$ for $0 \leqq i \leqq m$; and we infer from $zKe \leqq G$ that $G^{(i)} \cap Z^* = T \cap G \cap Z^* = zKe \cap G = zKe$. Since z belongs to Z, it belongs to $Z^* = Ke$ so that $z = ze$ belongs to zKe. It is a consequence of 13.1 and 13.2 that $G^{(i)}$ belongs to $D(G)$ and thus $G^{(i)}$ contains with z the subgroup $Z = zE(G)$ in $D(G)$, i.e.

$$Z \leqq G^{(i)} \cap Z^* = zKe \leqq G \cap Z^* = Z \quad \text{or} \quad Z = zKe.$$

Since the element c belongs to Z, and since $z = ze$, there exists an element c' such that $c = zec'e$.

Since U^* is an anti-cycle of order $n(Z)$ in $LA(K)$, it follows that $R(U^*) = eK$ is a cycle of order $n(Z)$ in $RA(K)$; and it follows from VI that $GeK = G^{(m-n(Z))}$ (remembering that G is a cycle of maximum order m in $RA(K)$). It is a consequence of Lemma 10.3 and of 13.3 that $GeKe_v = GeK \cap Ke_v = G^{(m-n)} \cap Z_v^* = G^{(m-n)} \cap G \cap Z_v^* = G^{(m-n)} \cap Z_v = \overline{Z}_v$ where $n = n(Z)$ and where $\overline{Z}_v$ is the uniquely determined subcycle of Z_v in $D(G)$ whose order is the minimum of $n(Z)$ and $n(Z_v)$. Since the order of the element c_v in Z_v does not exceed $n(Z)$, and since $Z = zKe$, it follows that c_v is an element in $\overline{Z}_v = GeKe_v = (G \cap Ke)eKe_v = (G \cap Z^*)eKe_v = ZeKe_v = zKeKe_v$; and hence there exist elements c_v', c_v'' such that $c_v = zec_v'ec_v''e_v$, since $z = ze$. Since almost all the c_v are null, we may assume that almost all the elements $ec_v'ec_v''e_v$ are null; and hence we may form the element $b = ec'e + \sum_v ec_v'ec_v''e_v$. Clearly $eb = b$ so that $Ub \leqq U^*b = Kfeb = 0$; and $zb = zec'e + \sum_v zec_v'ec_v''e_v = c + \sum_v c_v = t$. Thus we have shown that right-multiplication of G by the element b induces an $E(G)$-automorphism in $A(G)$ which maps U upon 0 and z upon t.

13.8. *Every finite*[54] *$E(G)$-automorphism of the abelian group G is contained in $A(G)$* (I; II, b; III; IV, 2; V; VI).

PROOF: If φ is a finite $E(G)$-automorphism of G, $N(\varphi)$ the set of all the elements in G that are mapped upon 0 by φ, then there exists by Lemma 6.1 a finite number of cycles $z_iE(G) = Z_i$ in $D(G)$ and a subgroup V in $D(G)$ such that G is the direct sum of V and the Z_i and such that $V \leqq N(\varphi)$, since G is by 13.5 a primary abelian operator group over the (primary) ring $E(G)$, and since $D(G) = D(G;E(G))$ by 13.1. Since $n(z_i^\varphi) \leqq n(Z_i)$, there exists by 13.7 an $E(G)$-automorphism φ_i in $A(G)$ which maps $V + \sum_{j \neq i} Z_j$ upon 0 and z_i upon z_i^φ. It is readily seen that $\varphi = \sum_i \varphi_i$; and consequently φ belongs to $A(G)$, since $A(G)$ is a ring of automorphisms.

13.9. *Every $E(G)$-automorphism of the abelian group G is contained in $A(G)$* (I; II, b; III; IV, 2; V; VI; VII).

---

[54] For definition and properties of finite automorphisms, cp. section **6.**

PROOF: If J is a right-annulet, then $J = R(L(J)) \leqq R(L(J) \cap G)$ and $L(J) \cap G \leqq L(R(L(J) \cap G))$. It is a consequence of II, b that $L(J)$ is the only left-annulet in K whose cross-cut with G is just $G \cap L(J)$ and thus we deduce from the last inequality that $L(J) \leqq L(R(L(J) \cap G))$. Hence $R(L(J) \cap G) = R(L(R(L(J) \cap G))) \leqq R(L(J)) = J \leqq R(L(J) \cap G)$ and thus we have shown that every right-annulet J in K satisfies:

$$J = R(L(J) \cap G).$$

To apply postulate VII we have to consider the class W of all the right-annulets in K which are cross-cuts of a finite number of anticycles in $RA(K)$. Clearly the right-annulet J belongs to the class W if, and only if, $L(J)$ is the join of a finite number of cycles in $LA(K)$; and it is a consequence of 13.1 and II, b that the latter property is equivalent to the condition: $L(J) \cap G$ is the sum of a finite number of cycles in $D(G)$.

If φ is an E-automorphism of G, J a right-annulet in the class W, then we denote by $f(J)$ the set of all the elements in K which induce by right-multiplication in G an automorphism ξ in $A(G)$, satisfying: $\xi \equiv \varphi \bmod \mathrm{P}(L(J) \cap G)$, i.e. $f(J)$ consists of all the elements x in K which satisfy: $yx = y^{\varphi}$ for every element y in $L(J) \cap G$.

If J is in the class W, then $L(J) \cap G$ is the sum of a finite number of cycles in $D(G) = D(G;E(G))$ and the right-annulet $\mathrm{P}(L(J) \cap G)$ in the automorphism ring $\mathrm{A}(G;E(G))$ of the primary abelian operator group G over the (primary) ring $E(G)$ (cf. 13.5) is the cross-cut of a finite number of anti-cycles in the partially ordered set of all the right-annulats in $\mathrm{A}(G;E(G))$, as may be inferred from the fundamental theorem of section **2**. Hence it follows from Theorem 6.3, (b) that there exists a finite $E(G)$-automorphism ρ of G satisfying: $\rho \equiv \varphi \bmod \mathrm{P}(L(J) \cap G)$; and this shows by 13.8 that $f(J)$ is not empty.

If the right-annulet H belongs to the class W, and if $H \leqq J$, then $L(J) \cap G \leqq L(H) \cap G$. If furthermore x is an element in $L(J) \cap G$, y an element in $f(J)$ and z an element in $f(H)$, then it follows that $xy = x^{\varphi} = xz$; and this shows that $y - z$ belongs to $R(L(J) \cap G) = J$. Now it is readily seen that $f(J)$ is a coset of K modulo J, and that $H \leqq J$ implies $f(H) \leqq f(J)$, i.e. $f(J)$ is a W-function. Hence there exists by VII one and only one element f in K such that $f(J) = J + f$ for every J in the class W.

If g is any element in G, then there exists by II, b one and only one left-annulet L_g such that $gE(G) = L_g \cap G$. Clearly $J_g = R(L_g)$ belongs to the class W, since $L(J_g) = L(R(L_g)) = L_g$ is a cycle in $LA(K)$. Since g belongs to $L(J_g) \cap G$, and since f belongs to $f(J_g)$, it follows that $gf = g^{\varphi}$; and this shows that φ belongs to $A(G)$.

It is a consequence of 13.5 that G is a primary abelian operator group over the (primary) ring $E(G)$ (with $m(E(G)) = m$); it is a consequence of 13.1 that $D(G;E(G)) = D(G)$ contains at least two independent cycles of maximum order $m = m(E(G))$; it is a consequence of 13.6 that K and $A(G)$ are essentially the same; and it is a consequence of 13.9 that $A(G)$ is exactly the ring of all the $E(G)$-automorphisms of G; and this completes the proof of our theorem.

14. Rings with the same Ideal Theory

The rings K and K^* are said to have *the same ideal theory*, if there exists a transformation f which maps in a biunivoque fashion the set of right-ideals in K upon the set of right-ideals in K^*, the set of left-ideals in K upon the set of left-ideals in K^*, the set of cross-cuts of right- and left-ideals in K upon the set of cross-cuts of right- and left-ideals in K^* and which satisfies the following rules:

(1) If R is a right-ideal in K and L a left-ideal in K, then $(R \cap L)^\varphi = R^\varphi \cap L^\varphi$; and $(LR)^\varphi = L^\varphi R^\varphi$.

(2) If S and T are both right-ideals or both left-ideals in K, then $(ST)^\varphi = S^\varphi T^\varphi$; and S is part of T if, and only if, S^φ is part of T^φ.

THEOREM: *Suppose that G is a primary abelian operator group over the (primary) ring E, that $D(G;E)$ contains at least three independent cycles of maximum order $m(E)$, and that the ring K contains an identity element* 1 *and satisfies postulate VII of section* **12**. *Then the automorphism ring* $\mathfrak{A}(G;E)$ *and the ring K have the same ideal theory if, and only if, they are isomorphic.*

PROOF: Isomorphic rings clearly have the same ideal theory. Thus let us suppose that the rings $\mathfrak{A}(G;E)$ and K have the same ideal theory. It is a consequence of the Existence Theorem of section **13** that the postulates I; II, b; III; IV, 3; V; VI are satisfied by the automorphism ring $\mathfrak{A}(G;E)$; and consequently the postulates I; II, b; III; IV, 3; V; VI; VII are satisfied by the ring K. Thus we infer from the Existence Theorem of section **13** the existence of a primary abelian operator group G^* over the (primary) ring E^* such that $m(E^*)$ is the maximum order of the cycles in $D(G^*;E^*)$ and such that the rings K and $\mathfrak{A}(G^*;E^*)$ are isomorphic. It is a consequence of the Fundamental Theorem of section **2** that there exists a projectivity of $D(G;E)$ upon the partially ordered set $LA(\mathfrak{A}(G;E))$ of the left-annulets in $\mathfrak{A}(G;E)$ and that there exists a projectivity of $LA(\mathfrak{A}(G^*;E^*))$ upon $D(G^*;E^*)$; and there exists a projectivity of $LA(\mathfrak{A}(G;E))$ upon $LA(\mathfrak{A}(G^*;E^*))$, since $\mathfrak{A}(G;E)$ and $\mathfrak{A}(G^*;E^*)$ have the same ideal theory. Thus there exists a projectivity of $D(G;E)$ upon $D(G^*;E^*)$; and it follows from **1G** that there exists an isomorphism η of G upon G^* and an isomorphism η of E upon E^* such that $(ge)^\eta = g^\eta e^\eta$ for g in G and e in E. Such an isomorphism η clearly induces an isomorphism of $\mathfrak{A}(G;E)$ upon the ring $\mathfrak{A}(G^*;E^*)$ isomorphic to K, and this completes the proof.

UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

BIBLIOGRAPHY

A. A. ALBERT.

(1) *Structure of Algebras*. A. M. S. Colloquium Publications **24**; New York, 1939.

K. ASANO.

(1) *Über verallgemeinerte Abelsche Gruppen mit hyperkomplexem Operatorenring und ihre Anwendungen*. Jap. Journ. of Math. **15** (1939), 231–253.

R. BAER.

(1) *The decomposition of abelian groups into direct summands*. Quart. Journ. of Math. **6** (1935), 222–232.

(2) *The significance of the system of subgroups for the structure of the group*. Amer. Journ. of Math. **61** (1938), 1–44.

(3) *Duality and commutativity of groups.* Duke Math. Journ. **5** (1940), 824–838.
(4) *A unified theory of projective spaces and finite abelian groups.* Transactions Amer. Math. Soc. **52** (1942), 283–343.

R. BRAUER.
(1) *Über Systeme hyperkomplexer Zahlen.* Math. Zeitschr. **29** (1929), 79–107.

H. FITTING.
(1) *Die Theorie der Automorphismenringe Abelscher Gruppen und ihr Analogon bei nichtkommutativen Gruppen.* Math. Ann. **107** (1932), 514–542.

G. KÖTHE.
(1) *Verallgemeinerte Abelsche Gruppen mit hyperkomplexem Operatorenring.* Math. Zeitschr. **39** (1934), 31–44.

P. E. LEWIS.
(1) *Characters of abelian groups.* Amer. Journ. of Math. **64** (1942), 81–105.

T. NAKAYAMA.
(1) *On Frobeniusean Algebras.* Ann. of Math. **40** (1939), 611–633; **42** (1941), 1–21.

J. VON NEUMANN.
(1) *Continuous Geometry.* Princeton Lectures 1936/37.

L. PONTRJAGIN.
(1) *Topological Groups.* Princeton 1939.

M. SHIFFMAN.
(1) *The ring of automorphisms of an abelian group.* Duke Journ. of Math. **5** (1939), 579–597.

K. SHODA.
(1) *Über den Automorphismenring bezw. die Automorphismengruppe einer endlichen Abelschen Gruppe.* Proc. Imp. Acad. Japan, **6** (1930), 10.
(2) *Über die Automorphismen einer endlichen Abelschen Gruppe.* Math. Ann. **100** (1928), 674–686.

TH. SKOLEM.
(1) *Zur Theorie der assoziativen Zahlensysteme.* Norske Videnskab Selskapet Skrifter Oslo, 1927, (12).

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

THEOREMS ON LINEAR COMBINATORIAL TOPOLOGY AND GENERAL MEASURE

BY R. RADO

(Received November 9, 1942)

1. Introduction

1.1 Consider a cube whose vertices are $a_1, a_2, \cdots, a_8$. Give to every one of its 12 edges some arbitrary orientation, denoting by $a_k \to a_l$ the fact that a_k and a_l are joined by an edge and that, in our orientation, a_k is its first and a_l its second endpoint. We may, for instance, agree to write $a_k \to a_l$ whenever a_k is nearer to a_1 than a_l.

Associate with every vertex a_n an integer $f(a_n) \geqq 0$. We shall be concerned with certain transformations T which operate on such functions f and which may be described as having the effect of "pushing a unit along some edge $a_k \to a_l$". More accurately, the transformation $T^{a_k,a_l,1}$ transforms f into f_1 where

$$\begin{aligned} f_1(a_k) &= f(a_k) - 1 \\ f_1(a_l) &= f(a_l) + 1 \\ f_1(a_n) &= f(a_n) \qquad (n \neq k, l). \end{aligned}$$

$T^{a_k,a_l,1}$ is only applicable to f provided $f(a_k) > 0$.

The object of this paper is to investigate, under similar but far more general circumstances, whether it is possible to transform by means of a finite number of suitable T-transformations, a given function $f(a_n)$ into a given function $g(a_n)$. We shall also consider two classes of more specialized T-transformations namely

(i) those (denoted by T') for which $f_1(a_k) \geqq g(a_k)$
(ii) those (denoted by T'') for which $f_1(a_l) \leqq f(a_k)$.

Special cases of the general problem have been treated before (see 1.4).

1.2 Let $\mathfrak{A} = \{a_{n_1}, a_{n_2}, \cdots, a_{n_r}\}$[1] be a non-empty set of vertices which has the following property: *whenever* $a_k \to a_l \prec \mathfrak{A}$ *then* $a_k \prec \mathfrak{A}$. We call such sets of vertices *sections*. If we adopt, for instance, the convention about $a_k \to a_l$ mentioned in 1.1, and if $a_1 \to a_2$, $a_1 \to a_3$ then $\{a_1, a_2, a_3\}$ is a section. A moment's consideration shows that, for every section $\mathfrak{A}$, the function f_1 of 1.1 satisfies

$$f_1(a_{n_1}) + f_1(a_{n_2}) + \cdots + f_1(a_{n_r}) \leqq f(a_{n_1}) + \cdots + f(a_{n_r}).$$

[1] Brackets { } are only employed to denote sets consisting of the elements specified between them. Relations $X \prec \mathfrak{S} \subset \mathfrak{S}'$ mean that X is element of the set $\mathfrak{S}$ which, in turn, is subset of the set $\mathfrak{S}'$. Sums, differences and products of sets have the meaning usually assigned to them in the theory of point sets.

Hence a *necessary* condition for the possibility of transforming, by means of suitable T-transformations, f into g is that, for every section $\mathfrak{A}$,

$$\sum_{a_n < \mathfrak{A}} f(a_n) \geqq \sum_{a_n < \mathfrak{A}} g(a_n). \tag{1}$$

Our main result states that, under more general circumstances which are explained in 1.3, these conditions are also *sufficient.*

1.3 I shall briefly indicate the kind of generalization to be made. Instead of the system of edges of a cube we shall consider, in the language of combinatorial topology, an arbitrary orientated one-dimensional complex or, using a different terminology, an arbitrary finite oriented graph Γ[2]. The functional values $f(x)$ attached to the "points" x of Γ, will not be non-negative integers but certain sets, subsets of a given abstract set S. The effect of a transformation $T^{a,b,A}$ upon f will be the removal of a subset A from the set $f(a)$ and the addition of the same set A to the set $f(b)$. Here, as before, $a \rightarrow b$ is an edge of Γ. Instead of postulating that after a number of suitable T-transformations f is changed into g we only require that f is transformed into a function f which is *equimeasurable* to the given function g, a term I am going to explain now.

The sets $f(x)$, $g(x)$ belong to a system of sets $A, B, \cdots$ for which, according to the definitions[3] in **1,** p. 62–64, a monotonic, additive and distributive measure $|A|$ is defined. The relevant definitions and facts are set out in section **2** of this note. All commonly employed types of measure are instances of this general measure definition. The system of sets $\bar{f}(x)$ and the system of sets $g(x)$ are called *equimeasurable* if, from the point of view of the measure in question, they are indistinguishable, i.e.: Whenever $\bar{A}$ is a set obtained from the sets $\bar{f}(x)$ by repeated applications of the processes of addition, subtraction and multiplication, and B is the set obtained by the same processes applied to the corresponding sets $g(x)$ then always $|A| = |B|$. The simple problem of 1.1 is a special case of this more general problem. For suppose that $f(a_n)$, $g(a_n)$ are non-negative integers. Let $f'(a_n)$, $g'(a_n)$ be sets containing exactly $f(a_n)$ and $g(a_n)$ elements respectively, and suppose that for $m \neq n$ the sets $f'(a_m)$, $f'(a_n)$ as well as the sets $g'(a_m)$, $g'(a_n)$ have no element in common. As measure $|A|$ we employ the number of elements of A. Then the generalized problem for the functions f', g' coincides with the special problem for f, g.

Returning to the general case, it will be obvious from fundamental properties of the measure that a *necessary* condition for the possibility of transforming $f(x)$ in the described way into a function $\bar{f}(x)$ which is equimeasurable with $g(x)$ is, that the measure of the set

$$\sum_{x < \mathfrak{A}} f(x)$$

---

[2] The main result of this note has been extended to the case of an arbitrary *infinite* graph. It is hoped that this case will be dealt with in a later paper.

[3] Numbers in heavy type refer to the bibliography at the end.

is not less than the measure of

$$\sum_{x \prec \mathfrak{A}} g(x).$$

Here $\mathfrak{A}$ is any section of Γ. It will be proved that these conditions are also *sufficient*, provided the given functions $f(x)$, $g(x)$ are such that sets $f(x)$ belonging to different points x have no element in common, and a similar condition holds for g. The case of arbitrary functions f, g will also be settled. It will be reduced to the special case just described, by familiar methods belonging to Boolean Algebra which have already been used in **1,** p. 79 section 2.

1.4 Theorem 5 below contains as special case the main result of **1** (Theorem I, p. 66). A simple instance where Theorem 4 may be employed is the proof of Muirhead's theorem on inequalities between symmetrical polynomials (**2,** p. 44 ff., in particular p. 46, (2). Also p. 63). Here the graph is

$$a_1 \rightarrow a_2 \rightarrow \cdots \rightarrow a_n ,$$

and the functional values $f(a_\nu)$ are non-negative numbers.

In a concluding section it is shown that the axioms imposed upon the measure are, in fact, necessary for the validity of our theorems. Furthermore, it is proved that, in the case of any graph and real numbers as functional values of f and g, the conditions (1) are, apart from certain trivial exceptions, independent from each other.

2. Definitions

2.1 Let Γ be a finite orientated graph, i.e. a finite, non-empty set of objects called *points*, here always denoted by the letters a, b, c, x, y (and a' etc.), together with a relation "$\rightarrow$" valid between certain ordered pairs of distinct points. $a \rightarrow b$ defines an "edge" of Γ.[4] The truth of $a \rightarrow b$ does not exclude the truth of $b \rightarrow a$, nor does necessarily every point x occur in some relation $a \rightarrow x$ or $x \rightarrow b$. Indeed, Γ may consist of points only, without any edges joining them. As already mentioned, $x \rightarrow x$ does not hold for any x.

Gothic capitals $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{G}$ denote sets of points of Γ. In particular $\mathfrak{G}$ is the set of all points,

$$\mathfrak{G} = \sum_x \{x\}.$$

Unless the contrary is stated, the letter $\mathfrak{A}$ (and $\mathfrak{A}'$ etc.) is used exclusively to denote sets which possess a certain characteristic property and which we call *sections* of Γ. A section is, by definition, a set $\mathfrak{A}$ which is such that *whenever*

$$a \rightarrow b \prec \mathfrak{A}$$

then

$$a \prec \mathfrak{A}.$$

For instance $\mathfrak{G}$ is a section, and so is the empty set 0.

[4] Other terms found in the literature are: node, knot, vertex for point, and branch for edge.

2.2 Let S be an abstract aggregate, fixed throughout. The letters A, B, C denote subsets of S. A measure $|A|$ is defined in S, in the sense of **1,** 62. This means that certain subsets of S are termed *measurable* sets. From now onwards A, B, C always denote measurable sets. We suppose that for every A, B the sets $A + B$; AB; $A - AB$ are measurable. The system of all measurable sets is divided into non-overlapping classes, and the class to which A belongs is denoted by $|A|$ and is called "the measure of A." The system of all measures $|A|$ is ordered by means of a transitive relation "$<$." Given any A, B exactly one of the three relations

$$|A| < |B|; \qquad |A| = |B|; \qquad |A| > |B|$$

holds. $|A| \leqq |B|$ denotes the logical sum of the first two of these relations. The measure is supposed to have the following properties.

Property (M): it is *monotonic*, i.e. $A \subset B$ implies $|A| \leqq |B|$.

Property (A): it is *additive*, i.e.

$$|A_1| = |B_1|; \qquad |A_2| = |B_2|; \qquad A_1A_2 = B_1B_2 = 0^5$$

imply

$$|A_1 + A_2| = |B_1 + B_2|.$$

Property (D): it is *distributive*, i.e.

$$|A| \geqq |B_1 + B_2|; \qquad B_1B_2 = 0$$

imply the existence of sets A_1, A_2 satisfying

$$A_1 + A_2 \subset A; \qquad A_1A_2 = 0;$$

$$|A_1| = |B_1|; \qquad |A_2| = |B_2|.$$

In particular it follows from (D) (put $B_2 = 0$) that

$$|A| \geqq |B_1|$$

implies the existence of $A_1 \subset A$ such that

$$|A_1| = |B_1|.$$

For convenience of reference let us call this property of the measure *Property* (C) (the measure is *continuous*).

In fact, (D) is not used explicitly in this paper, everywhere the special case (C) is sufficient. The only occasion where (D) is needed is the proof of Lemma 1, and this proof is not given here since Lemma 1 is taken from **1.** (D) is, however, necessary for Theorem 2 to be true as will be shown in 8.1.

The set A is said to be of *finite measure* (notation: $|A| < \infty$) if $A' \subset A$; $|A'| = |A|$ imply $|A - A'| = 0$.

Instances where this theory of measure applies are:

(i) S is any abstract aggregate, measurable for all subsets of S, and we write $|A| < |B|$ if, and only if, the power of A, as defined in the theory of aggre-

[5] By 0 we denote, without risking any confusion, the number zero as well as the empty set and the measure of the empty set.

gates, is less than the power of B. Here sets of finite measure in our sense are all finite subsets of S.

(ii) S is a euclidean space, its measurable subsets are all Lebesgue-measurable sets in S, and $|A| < |B|$ signifies that the Lebesgue measure of A is less than that of B. Here sets of finite measure in our sense are those of finite Lebesgue measure.

(iii) A more sophisticated example: S is the set of all real numbers, its measurable subsets are all sets A of the form

$$A = \sum_{\nu=1}^{n} [\alpha_\nu, \beta_\nu)$$

where $[\alpha, \beta)$ is the set of all t satisfying $\alpha \leqq t < \beta$, i.e. all sets consisting of a finite number of intervals which are closed on the left and open on the right. Define a number $r(A)$ as follows. If A contains no integer then let $h(A)$ be the (necessarily finite) number of numbers in A which are of the form $k + \frac{1}{l}$ where k and l are integers, $l > 1$. Then put

$$r(A) = \frac{h(A)}{h(A) + 1}.$$

If A contains exactly $m(> 0)$ integers then put

$$r(A) = m.$$

Define a measure by writing $|A| < |B|$ if, and only if $r(A) < r(B)$. It is easily verified that (M), (A), (D) hold. Sets of finite measure are those which contain no integers. If for instance

$$A = [0, 4); \qquad B = [4\tfrac{1}{2}, 5); \qquad C = [5, 6)$$

then

$$0 < |B| < |C| = |B + C| < |A| = |A + B| < |A + C| = |A + B + C|; \qquad AB = AC = BC = 0.$$

2.3 Let $f(x)$, $g(x)$ be functions which are defined for every point x, and whose functional values are measurable subsets of S. Throughout this note g will be kept fixed. It is supposed to satisfy

(2) $$|g(x)| < \infty \qquad \textit{for all } x.$$

g is called *decreasing* if $a \to b$ implies $|g(a)| \geqq |g(b)|$. $f(x)$ is quite arbitrary. It will be subjected to certain transformations T. Choose any edge $a \to b$ and any set $A \subset f(x)$. Then $T^{a,b,A}$ denotes a transformation which transforms $f(x)$ into

(3) $$T^{a,b,A} f = f_1$$

defined by:

$$f_1(a) = f(a) - A; \qquad f_1(b) = f(b) + A; \qquad f_1(x) = f(x) \quad \textit{for} \quad x \neq a, b.$$

If, in addition, $|f_1(a)| \geqq |g(a)|$ then (3) is called a *T'-transformation.* If $|f_1(b)| \leqq |f(a)|$ we call (3) a *T''-transformation.* A function $\tilde{f}$ is called a *T-transform* of f if $\tilde{f}$ can be obtained from f by applying a finite number of suitable transformations $T^{a,b,A}$. Analogous meanings are attached to the terms *T'-transform, T''-transform.*

Finally, we introduce the abbreviation

$$(f, \mathfrak{B}) = \sum_{x \prec \mathfrak{B}} f(x).^{6}$$

3. Statement of the theorems

3.1 Common hypotheses for theorems 1–4 are (2) and

$$f(x)f(y) = g(x)g(y) = 0 \qquad \text{for } x \neq y. \tag{4}$$

THEOREM 1. *Given $f(x)$ and $g(x)$ necessary and sufficient for f to have a T'-transform $\tilde{f}$ satisfying*

$$|\tilde{f}(x)| \geqq |g(x)| \qquad \text{for all } x, \tag{5}$$

is that

$$|(f, \mathfrak{A})| \geqq |(g, \mathfrak{A})| \qquad \text{for all } \mathfrak{A}. \tag{6}$$

THEOREM 2. *Given $f(x)$ and $g(x)$, necessary and sufficient for f to have a T'-transform $\tilde{f}$ satisfying*

$$|\tilde{f}(x)| = |g(x)| \qquad \text{for all } x, \tag{7}$$

is that, in addition to (6),

$$|(f, \mathfrak{A})| = |(g, \mathfrak{A})|. \tag{8}$$

THEOREM 3. *Given $f(x)$ and $g(x)$, and supposing that $g(x)$ is decreasing, necessary and sufficient for f to have a T''-transform $\tilde{f}$ satisfying* (5) *is that* (6) *should be true.*

THEOREM 4. *Given $f(x)$ and $g(x)$, and supposing that $g(x)$ is decreasing, necessary and sufficient for f to have a T''-transform $\tilde{f}$ satisfying* (7) *is that* (6) *and* (8) *should be true.*

3.2 Before passing to the general case in which (4) is not necessarily satisfied we want to introduce some more notations. A relation

$$a \longrightarrow\!\!\!\!\!\rightarrow b$$

means, by definition, that there are points $x_0, x_1, \cdots, x_n$ such that

$$a = x_0 \rightarrow x_1 \rightarrow \cdots \rightarrow x_n = b.$$

Here $n \geqq 0$. In particular, $a \longrightarrow\!\!\!\!\!\rightarrow a$ holds for every a, and

$$a \longrightarrow\!\!\!\!\!\rightarrow b \longrightarrow\!\!\!\!\!\rightarrow c$$

implies $a \longrightarrow\!\!\!\!\!\rightarrow c$.

---

[6] Empty sums have the value 0.

Let $[\mathfrak{B}]$[7] denote the smallest section of Γ which contains $\mathfrak{B}$, i.e.

$$[\mathfrak{B}] = \prod_{\mathfrak{B} \subset \mathfrak{A}} \mathfrak{A} = \sum_{x \to\to y \prec \mathfrak{B}} \{x\}.$$

We have, for instance,

$$\mathfrak{A}' = [\mathfrak{A}'] = \sum_{x \prec \mathfrak{A}'} [\{x\}] \tag{9}$$

for every $\mathfrak{A}'$.

If $\mathfrak{B} \neq 0$ we denote by $F(\mathfrak{B})$ the set consisting of those elements of S which belong to every set $f(x)$ when $x \prec \mathfrak{B}$, and to none of the remaining sets $f(x)$, i.e.

$$F(\mathfrak{B}) = \prod_{x \prec \mathfrak{B}} f(x) \prod_{y \text{ not in } \mathfrak{B}} (S - f(y)).^{8} \tag{10}$$

Put $F(0) = 0$. Similarly $\bar{F}(\mathfrak{B})$, $G(\mathfrak{B})$ etc. are defined in terms of $\bar{f}(x)$, $g(x)$ etc. Obviously[9]

$$F(\mathfrak{B}_1)F(\mathfrak{B}_2) = 0 \quad \text{for} \quad \mathfrak{B}_1 \neq \mathfrak{B}_2 .$$

Every set which is obtained from the sets $f(x)$ by the processes of addition, subtraction and multiplication, is a sum of certain sets $F(\mathfrak{B})$. This, in conjunction with property (A), clearly shows that *$f(x)$ and $g(x)$ are equimeasurable* (see 1.2) *if, and only if,*

$$|F(\mathfrak{B})| = |G(\mathfrak{B})| \qquad \textit{for every } \mathfrak{B}. \tag{11}$$

Let $\psi(x)$ be a notation for any function which associates with every point x a point $\psi(x)$ in such a way that

$$\psi(x) \to\to x \qquad \textit{for all } x.$$

$\psi_{\mathfrak{B}}(x)$ is a notation for those among the $\psi(x)$ which satisfy, in addition,

$$\psi_{\mathfrak{B}}(x_1) \neq \psi_{\mathfrak{B}}(x_2)$$

whenever x_1, x_2 are distinct elements of $\mathfrak{B}$, e.g. $\psi_{\mathfrak{B}}(x) = x$ is such a function. If, in particular, $\mathfrak{B} = 0$ or $\mathfrak{B}$ consists of one single element then every function $\psi(x)$ is at the same time a function $\psi_{\mathfrak{B}}(x)$.

Now put, for $\mathfrak{B} \neq 0$,

$$F^*(\mathfrak{B}) = \Big(\sum_{\psi_{\mathfrak{B}}} \prod_{x \prec \mathfrak{B}} f(\psi_{\mathfrak{B}}(x))\Big) \prod_{y \text{ not in } \mathfrak{B}} (S - f(y)) \tag{12}$$

and

$$F^*(0) = 0.$$

F^*, G^* etc. are defined in terms of $\bar{f}$, g etc. as F^* was defined in terms of f.

In theorems 5 and 6 $f(x)$ and $g(x)$ are arbitrary functions, not necessarily satisfying (4). But (2) is true throughout.

---

[7] Square brackets will only be used in this special sense.

[8] $y \not\prec \mathfrak{B}$ is the negation of $y \prec \mathfrak{B}$, empty products have the value S.

[9] A formal proof is given in Footnote [7,2].

THEOREM 5. *Given $f(x)$ and $g(x)$, necessary and sufficient for f to have a T-transform $\bar{f}$ satisfying*

$$|\bar{F}(\mathfrak{B})| \geqq |G(\mathfrak{B})| \qquad \text{for all } \mathfrak{B} \tag{13}$$

is that

$$\left|\sum_{\kappa=1}^{k} F^*(\mathfrak{B}_\kappa)\right| \geqq \left|\sum_{\kappa=1}^{k} G^*(\mathfrak{B}_\kappa)\right| \tag{14}$$

for every $k > 0$ and every selection of k sets $\mathfrak{B}_\kappa$.

THEOREM 6. *Given $f(x)$ and $g(x)$, necessary and sufficient for f to have a T-transform $\bar{f}$ which is equimeasurable to g, i.e. which satisfies*

$$|\bar{F}(\mathfrak{B})| = |G(\mathfrak{B})| \qquad \text{for all } \mathfrak{B} \tag{15}$$

is that (8) *and* (14) *should be true.*

Evidently (15) and (7) are equivalent statements if (4) holds.

3.3 Let us consider some special cases of theorems 5 and 6. The necessary verifications are easily made and are left to the reader.

(i) $\mathfrak{G} = \{1, 2, \cdots, n, n+1\}$. The edges of Γ are

$$\nu \to n+1 \qquad (1 \leqq \nu \leqq n).$$

If $\mathfrak{B} \subset \{1, 2, \cdots, n\}$ then

$$F^*(\mathfrak{B}) = F(\mathfrak{B}),$$

$$F^*(\mathfrak{B} + \{n+1\}) = \Big(\prod_{x \prec \mathfrak{B}} f(x)\Big)\Big(\sum_{y \text{ not in } \mathfrak{B}} f(y)\Big).$$

Now suppose that, in particular,

$$f(n+1) = \sum_{\nu=1}^{n} f(\nu); \qquad g(n+1) = \sum_{\nu=1}^{n} g(\nu),$$

while

$$f(1), \cdots, f(n), \qquad g(1), \cdots, g(n)$$

are $2n$ arbitrary sets the $g(\nu)$ $(1 \leqq \nu \leqq n)$, however, of finite measure[10]. Then the only effect of a transformation T upon the function $f(x)$ is the removal of some set A from one of the sets $f(\nu)$ $(1 \leqq \nu \leqq n)$, and vice versa, every such removal is effected by a transformation T. Furthermore, for every $\mathfrak{B} \subset \{1, \cdots, n\}$, we have

$$F^*(\mathfrak{B}) = 0,$$

$$F^*(\mathfrak{B} + \{n+1\}) = \prod_{x \prec \mathfrak{B}} f(x) \qquad (\text{if } \mathfrak{B} \neq 0),$$

$$F^*(\{n+1\}) = f(n+1),$$

[10] By Lemma 1 below, $g(1) + \cdots + g(n)$ is of finite measure.

and analogous equations with g instead of f. Hence Theorem 6 reduces to the main result of **1** (Theorem I) which states:

Given $2n$ sets $A_1, \cdots, A_n, B_1, \cdots, B_n$, of which the B_ν are of finite measure, it is possible to select subsets $\bar{A}_\nu \subset A_\nu$ in such a way that the system $\bar{A}_1, \cdots, \bar{A}_n$ is equimeasurable to the system $B_1, \cdots, B_n$, provided the following (necessary and sufficient) condition holds: For every set A obtained from the A_ν by any applications of the processes of addition and multiplication, and for the set B obtained by applying the same processes to the corresponding sets B_ν, we have $|A| \geqq |B|$.

(ii) $\mathfrak{G} = \{1, 2, \cdots, n\}$. The edges of Γ are

$$1 \to 2 \to 3 \to \cdots \to n \to 1.$$

If $k > 0$; $1 \leqq \nu_1 < \nu_2 < \cdots < \nu_k \leqq n$ then

$$F^*(\{\nu_1, \cdots, \nu_k\}) = \sum_{1 \leqq \lambda_1 < \cdots < \lambda_k \leqq n} f(\lambda_1) f(\lambda_2) \cdots f(\lambda_k).$$

Hence the result: *Given sets $A_1, \cdots, A_n, B_1, \cdots, B_n$ ($|B_\nu| < \infty$) it is possible, by removing suitable subsets from some A_ν's and adding them to the corresponding set $A_{\nu+1}$ ($A_{n+1} = A_1$) to change $A_1, \cdots, A_n$ into a system equimeasurable to $B_1, \cdots, B_n$ provided the following (necessary and sufficient) conditions hold:*

$$(16) \quad \begin{cases} |A_1 + A_2 + \cdots + A_n| = |B_1 + \cdots + B_n| \\ |A_1A_2 + A_1A_3 + \cdots + A_{n-1}A_n| \geqq |B_1B_2 + \cdots + B_{n-1}B_n| \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ |A_1A_2 \cdots A_n| \geqq |B_1B_2 \cdots B_n|. \end{cases}$$

The kth relation (16) involves all $\binom{n}{k}$ products of k sets A_ν and B_ν respectively.

In fact, (16) is the solubility-condition for any graph with n points which satisfies $x \longrightarrow y$ for every x, y.

(iii) Γ is arbitrary, f and g satisfy (4). Then

$$F(\{a\}) = f(a),$$

$$F^*(\{a\}) = \sum_{x \to\to a} f(x) = (f, [\{a\}]),$$

$$F(\mathfrak{B}) = F^*(\mathfrak{B}) = 0$$

if $\mathfrak{B}$ contains more than one element, (14) becomes

$$\left| \sum_{\kappa=1}^{k} (f, [\{a_\kappa\}]) \right| \geqq \left| \sum_{\kappa=1}^{k} (g, [\{a_\kappa\}]) \right|,$$

where $k > 0$ and the a_κ are arbitrary. By (9) this is the same as

$$|(f, \mathfrak{A})| \geqq |(g, \mathfrak{A})|$$

where

$$\mathfrak{A} = \sum_{\kappa=1}^{k} [\{a_\kappa\}]$$

is any arbitrary section. Hence Theorem 5 reduces to Theorem 1 if we weaken the latter's assertion by writing T for T''.

4. Preliminary remarks on theorems 1–4

4.1 Property (A) of the measure shows that, whenever

$$|A_\nu| = |B_\nu| \qquad (1 \leqq \nu \leqq 3),$$
$$A_\mu A_\nu = B_\mu B_\nu = 0 \qquad (1 \leqq \mu < \nu \leqq 3)$$

we have

$$|A_1 + A_2| = |B_1 + B_2|,$$
$$|A_1 + A_2 + A_3| = |(A_1 + A_2) + A_3| = |(B_1 + B_2) + B_3| = |B_1 + B_2 + B_3|.$$

Similarly for sums of more than three terms.

All facts on sets of finite measure which we shall have to use are contained in

LEMMA 1. (i) *If* $|A| \leqq |B| < \infty$ *then* $|A| < \infty$.

(ii) *If* $|A| < \infty$; $|B| < \infty$, *then* $|A + B| < \infty$.

(i) is Lemma 5, (ii) is Lemma 7 of **1.**

4.2 LEMMA 2. *If* $n \geqq 2$,

$$|A_\nu| \leqq |B_\nu| \qquad (1 \leqq \nu \leqq n), \tag{17}$$
$$A_\mu A_\nu = B_\mu B_\nu = 0 \qquad (1 \leqq \mu < \nu \leqq n)$$

then

$$|A_1 + \cdots + A_n| \leqq |B_1 + \cdots + B_n|. \tag{18}$$

If $|A_\nu| < \infty$ *for all* ν, *and if there is equality in* (18) *then there is equality in* (17) *for every* ν.

PROOF. By (17) and (C) there are sets $B'_\nu \subset B_\nu$ such that

$$|A_\nu| = |B'_\nu| \qquad (1 \leqq \nu \leqq n).$$

Then, by (A), in the form 4.1, and (M),

$$\left|\sum_1^n A_\nu\right| = \left|\sum_1^n B'_\nu\right| \leqq \left|B_1 + \sum_2^n B'_\nu\right| \leqq \left|\sum_1^n B_\nu\right|, \tag{19}$$

and therefore (18). If $|A_\nu| < \infty$ for all ν, and if there is equality in (18) then, by (19) and Lemma 1,

$$\sum_1^n B'_\nu \subset B_1 + \sum_2^n B'_\nu$$
$$\left|\sum_1^n B'_\nu\right| = \left|B_1 + \sum_2^n B'_\nu\right| \leqq \left|\sum_1^n B_\nu\right| = \left|\sum_1^n A_\nu\right| < \infty,$$
$$|B_1 - B'_1| = 0,$$
$$|A_1| = |B'_1| = |B'_1 + 0| = |B'_1 + (B_1 - B'_1)| \qquad (by\ (A))$$
$$= |B_1|.$$

Therefore

$$|A_\nu| = |B_\nu| \qquad (1 \leqq \nu \leqq n),$$

and the lemma is proved.

4.3 If $a \rightarrow b$, and if $\mathfrak{A}$ is any section the

either $$a, b \prec \mathfrak{A}$$

or $$a \prec \mathfrak{A}; \quad b \not\prec \mathfrak{A}$$

or $$a, b \not\prec \mathfrak{A}.$$

For the remaining possibility

$$a \not\prec \mathfrak{A}; \quad b \prec \mathfrak{A}$$

is excluded by the definition of sections. Now let

$$T^{a,b,A} f = f_1 .$$

Then it follows that

$$(f_1 , \mathfrak{A}) = \begin{cases} (f, \mathfrak{A}) - A & \text{if } a \prec \mathfrak{A}; b \not\prec \mathfrak{A} \\ (f, \mathfrak{A}) & \text{otherwise.} \end{cases}$$

Hence every T-transform $\tilde{f}$ of f satisfies

$$(\tilde{f}, \mathfrak{A}) \subset (f, \mathfrak{A}).$$

If, in addition, (5) holds then, by (M) and Lemma 2,

$$|(f, \mathfrak{A})| \geqq |(\tilde{f}, \mathfrak{A})| \geqq |(g, \mathfrak{A})| .$$

Therefore (6) is a necessary condition in theorems 1–4.

Clearly, the same function $\tilde{f}$ satisfies

$$(\tilde{f}, \mathfrak{A}) = (f, \mathfrak{A}).$$

If, moreover, (7) is satisfied then

$$|(f, \mathfrak{A})| = |(\tilde{f}, \mathfrak{A})| = |(g, \mathfrak{A})| , \tag{20}$$

i.e. (8). We have thus shown that (6) and (8) are necessary conditions in theorems 2 and 4.

4.4 Theorems 2 and 4 are easily deduced from theorems 1 and 3 respectively. For suppose that (6) and (8) are satisfied, and that, (in the case of Theorem 4) $g(x)$ is decreasing. Then by Theorem 1 (or Theorem 3) there is a T'-transform (or a T'''-transform) $\tilde{f}$ of f for which (5) holds. Then, once more, (20) follows, and (7) is a consequence of Lemma 2. Here we must observe that it follows from (2), (20) and Lemma 1 that $|\tilde{f}(x)| < \infty$.

To sum up, all we have to show in order to complete the proofs of theorems 1–4 is that (6) is sufficient for the existence of a T'-transform and, in case g is decreasing, of a T'''-transform $\tilde{f}$ satisfying (5).

5. Proof of theorems 1 and 2

5.1 Throughout this section (2) and (4) are supposed to hold. We introduce some further definitions and notations. A relation

$$a \equiv b \tag{21}$$

is, by definition, equivalent to

$$a \longrightarrow b \longrightarrow a,$$

and $a \not\equiv b$ is the negation of (21). An edge $a \to b$ is called *singular* if $a \equiv b$, and *regular* if $a \not\equiv b$. Singular edges are those which occur in closed cycles

$$x_1 \to x_2 \to x_3 \to \cdots \to x_n \to x_1 .$$

Let $\rho(\Gamma)$ denote the number of regular edges of Γ. (21) defines an equivalence relation and hence a subdivision of $\mathfrak{A}$ into mutually exclusive classes. The class to which any point x_o belongs is denoted by

$$\mathfrak{K}(x_o) = \sum_{x \equiv x_o} \{x\}.$$

Finally, put

$$p(x_o, \Gamma) = \sum_{x \to\to x_o} 1,$$

$$q(f, x_o, \Gamma) = \sum_{|f(x)| \leqq |f(x_o)|} 1,$$

$$s(f, \Gamma) = \sum_x p(x, \Gamma) q(f, x, \Gamma).$$

p counts the number of "ancestors" of x_o while q describes the position of $|f(x_o)|$ relative to all measures $|f(x)|$.

5.2 LEMMA 3. *If* $m \geqq 1$; $|A_1| \leqq |B_1|$,

$$|A_\mu + A| > |B_\mu| \qquad (1 \leqq \mu \leqq m) \tag{22}$$

then there is $A' \subset A$ *and* μ_o $(1 \leqq \mu_o \leqq m)$ *such that*

$$|A_\mu + A'| \begin{cases} \geqq |B_\mu| & (1 \leqq \mu \leqq m) \\ = |B_\mu| & (\mu = \mu_o). \end{cases}$$

This is Lemma 2 of **1.** Its assertion remains, of course, valid if in (22) we allow "$\geqq$" instead of "$>$."

LEMMA 4. *Let*

$$a_o \to b_o ,$$

$$a_o \prec \mathfrak{A}_o ; \qquad b_o \not\prec \mathfrak{A}_o .$$

Suppose Γ' *is the graph which is obtained from* Γ *by removing the edge* $a_o \to b_o$, *i.e.*

(i) $\mathfrak{A}' = \mathfrak{A}$;

(ii) $x \to y$ *is true in* Γ' *if, and only if,* $x \to y$ *is true in* Γ *and* $(x, y) \neq (a_o, b_o)$. *Let* $\mathfrak{A}'$ *be a section of* Γ' *(but not necessarily a section of* Γ*). Then* $\mathfrak{A}_o + \mathfrak{A}'$ *and* $\mathfrak{A}_o\mathfrak{A}'$ *are sections of* Γ.

PROOF. In the following proof the symbol "$\to$" is only used to denote edges of Γ. The characteristic property of $\mathfrak{A}'$ is: *Whenever*

$$a \to b \prec \mathfrak{A}'; \qquad (a, b) \neq (a_o, b_o)$$

then

$$a \prec \mathfrak{A}'.$$

1) If $\mathfrak{A}_o + \mathfrak{A}'$ would not be a section of Γ then there would exist points a, b such that

$$a \to b \prec \mathfrak{A}_o + \mathfrak{A}'; \qquad a \not\prec \mathfrak{A}_o + \mathfrak{A}'.$$

Then we would conclude that

$$a \not\prec \mathfrak{A}_o; \qquad b \not\prec \mathfrak{A}_o; \qquad b \prec \mathfrak{A}'; \qquad a \neq a_o; \qquad a \prec \mathfrak{A}'$$

which is a contradiction.

2) If $\mathfrak{A}_o\mathfrak{A}'$ would not be a section of Γ then there would exist points a, b such that

$$a \to b \prec \mathfrak{A}_o\mathfrak{A}'; \qquad a \not\prec \mathfrak{A}_o\mathfrak{A}'.$$

Then we would conclude that

$$b \prec \mathfrak{A}_o; \qquad a \prec \mathfrak{A}_o; \qquad a \not\prec \mathfrak{A}'; \qquad b \neq b_o; \qquad b \not\prec \mathfrak{A}'$$

which is a contradiction.

5.3 LEMMA 5. *Notations and hypotheses as in Lemma* 4. *In addition we suppose that* (6) *holds and that*

$$|(f, \mathfrak{A}_o)| = |(g, \mathfrak{A}_o)|.$$

Then

$$|(f, \mathfrak{A}')| \geqq |(g, \mathfrak{A}')| \tag{23}$$

for all sections $\mathfrak{A}'$ *of* Γ'.

PROOF.

$$\mathfrak{A}' = \mathfrak{A}_o\mathfrak{A}' + ((\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o),$$

$$(f, \mathfrak{A}') = (f, \mathfrak{A}_o\mathfrak{A}') + (f, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o). \tag{24}$$

By Lemma 4 and (6),

$$|(f, \mathfrak{A}_o\mathfrak{A}')| \geqq |(g, \mathfrak{A}_o\mathfrak{A}')|, \tag{25}$$

$$|(f, \mathfrak{A}_o + \mathfrak{A}')| \geqq |(g, \mathfrak{A}_o + \mathfrak{A}')|. \tag{26}$$

If we assume that

$$|(f, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o)| < |(g, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o)| \tag{27}$$

we obtain a contradiction. For then, by (2) and Lemma 1,

$$|(f, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o)| < \infty,$$

$$|(f, \mathfrak{A}_o)| = |(g, \mathfrak{A}_o)| < \infty,$$

and hence, by Lemma 2,

$$|(f, \mathfrak{A}_o + \mathfrak{A}')| = |(f, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o) + (f, \mathfrak{A}_o)|$$

$$< |(g, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o) + (g, \mathfrak{A}_o)| = |(g, \mathfrak{A}_o + \mathfrak{A}')|$$

which contradicts (26). Hence (27) is not true, i.e.

$$|(f, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o)| \geqq |(g, (\mathfrak{A}_o + \mathfrak{A}') - \mathfrak{A}_o)|. \tag{28}$$

(23) follows from (24), (25), (28) and Lemma 2.

5.4 LEMMA 6. *Let* $\mathfrak{K}_o = \mathfrak{K}(x_o)$;

$$|(f, \mathfrak{K}_o)| \geqq |(g, \mathfrak{K}_o)|. \tag{29}$$

Then there exists a T′-transform $\bar{f}$ *of* f *satisfying*

$$|\bar{f}(x)| \geqq |g(x)| \qquad (x \prec \mathfrak{K}_o),$$

$$\bar{f}(x) = f(x) \qquad (x \not\prec \mathfrak{K}_o).$$

PROOF. Define an operator O_1 as follows. O_1 is applicable to every function $f(x)$ which satisfies (29) and, in addition,

$$|f(x_1)| < |g(x_1)| \tag{30}$$

for at least one $x_1 \prec \mathfrak{K}_o$. We may assume that O_1 is applicable to the given function f. Choose one point $x_1 \prec \mathfrak{K}_o$ which satisfies (30).[11] Then there is $x_2 \prec \mathfrak{K}_o$ such that

$$|f(x_2)| > |g(x_2)|.$$

For if we assume that

$$|f(x)| \leqq |g(x)| \quad \textit{for every} \quad x \prec \mathfrak{K}_o$$

then Lemma 2, in connection with (30), would lead to a contradiction against (29). It now follows from the definition of $\mathfrak{K}_o$ that there exists a chain of edges

$$x_2 = a_o \to a_1 \to \cdots \to a_n = x_1.$$

[11] We shall often have to define operations which necessitate a uniformly bounded number of selections of elements from sets whose elements are points or edges of sections. We make these operations definite by numbering, once for all, the points, edges and sections of Γ and choosing every time the first one. Selections made in virtue of (D) cannot in general be made definite.

Then $n > 0$. By definition of $\mathfrak{K}_o$ we have $x_1 \longrightarrow\!\!\!\!\!\rightarrow x_2$ and therefore

$$a_\nu < \mathfrak{K}_o \qquad (0 \leqq \nu \leqq n).$$

There is a largest index ν_o satisfying

$$0 \leqq \nu_o < n; \qquad |f(a_{\nu_o})| > |g(a_{\nu_o})|,$$

and then a smallest index ν_1 satisfying

$$\nu_o < \nu_1 \leqq n; \qquad |f(a_{\nu_1})| \neq |g(a_{\nu_1})|.$$

Then

$$|f(a_\nu)| \begin{cases} > |g(a_\nu)| & (\nu = \nu_o) \\ = |g(a_\nu)| & (\nu_o < \nu < \nu_1) \\ < |g(a_\nu)| & (\nu = \nu_1). \end{cases}$$

(i) If $\nu_o + 1 < \nu_1$ then there exists, by (C), a set $A_o \subset f(a_{\nu_o})$ such that

$$|f(a_{\nu_o}) - A_o| = |g(a_{\nu_o})|.$$

Then

$$T^{a_{\nu_o}, a_{\nu_o+1}, A_o} f = f_1,$$

is a T'-transform of f,

$$|f_1(a_{\nu_o})| = |g(a_{\nu_o})|,$$
$$|f_1(a_{\nu_o+1})| = |f(a_{\nu_o+1}) + A_o| \geqq |f(a_{\nu_o+1})|.$$

In the last relation equality is impossible. For if

$$|f(a_{\nu_o+1}) + A_o| = |f(a_{\nu_o+1})| = |g(a_{\nu_o+1})| < \infty$$

then, in view of

$$A_o f(a_{\nu_o+1}) \subset f(a_\nu) f(a_{\nu_o+1}) = 0,$$

we would conclude that $|A_o| = 0$ and therefore

$$|g(a_{\nu_o})| = |f(a_{\nu_o}) - A_o| = |f(a_{\nu_o})|$$

which is a contradiction. Hence

$$|f_1(a_\nu)| \begin{cases} > |g(a_\nu)| & (\nu = \nu_o + 1) \\ = |g(a_\nu)| & (\nu_o + 1 < \nu < \nu_1) \\ < |g(a_\nu)| & (\nu = \nu_1). \end{cases}$$

(ii) If $\nu_o + 2 < \nu_1$ then we treat f_1, $\nu_o + 1$ in the same way as f, ν_o were treated in (i), etc. After $\nu_1 - \nu_o - 1$ such T'-transformations (no transformation at all if $\nu_o + 1 = \nu_1$) we arrive at a T'-transform f' of f which satisfies

$$|f'(a_{\nu_1-1})| > |g(a_{\nu_1-1})|,$$
$$|f'(a_{\nu_1})| < |g(a_{\nu_1})|,$$
$$|f'(x)| = |f(x)| \qquad \text{for } x \neq a_{\nu_o}, a_{\nu_1-1}, a_{\nu_1}.$$

Also,

$$|f'(a_{\nu_0})| = |g(a_{\nu_0})| \qquad \text{if } \nu_0 + 1 < \nu_1 .$$

There is a set $A' \subset f'(a_{\nu_1-1})$ satisfying

$$|f'(a_{\nu_1-1}) - A'| = |g(a_{\nu_1-1})| .$$

Then

$$T^{a_{\nu_1-1}, a_{\nu_1}, A'} f' = f''$$

is a T'-transformation. We put, by definition,

$$O_1 f = f''$$

Then

$$|f''(a_\nu)| = |g(a_\nu)| \qquad (\nu_0 \leqq \nu < \nu_1),$$

$$f''(x) = f(x) \qquad (x \neq a_{\nu_0}, a_{\nu_0+1}, \cdots, a_{\nu_1}).$$

In particular, there are more points x satisfying

$$|f''(x)| = |g(x)|$$

than points x satisfying

$$|f(x)| = |g(x)| .$$

For every x of the second kind is also of the first kind, and

$$|f''(a_{\nu_0})| = |g(a_{\nu_0})| ; \qquad |f(a_{\nu_0})| > |g(a_{\nu_0})| .$$

Therefore O_1 can only be applied to f a bounded number of times in succession, and the result of the last possible application is a T'-transform $\tilde{f}$ of f which has the required properties.

5.5 LEMMA 7. *Let* $\mathfrak{K}_0 = \mathfrak{K}(x_0)$. *Suppose that*

$$|f(x)| \geqq |g(x)| \qquad (x \prec \mathfrak{K}_0). \tag{31}$$

Then there is a T'*-transform* $\tilde{f}$ *of* f *satisfying*

$$|\tilde{f}(x)| = |g(x)| \qquad (x \prec \mathfrak{K}_0 - \{x_0\}),$$

$$\tilde{f}(x) = f(x) \qquad (x \not\prec \mathfrak{K}_0).$$

PROOF. Define an operator O_2 as follows. O_2 is applicable to every f which satisfies (31) and, in addition,

$$|f(x_1)| > |g(x_1)|$$

for at least one point $x_1 \prec \mathfrak{K}_0 - \{x_0\}$. We may assume that O_2 is applicable to the given function f. Choose one of the above-mentioned points x_1. Then there is a chain

$$x_1 = a_0 \to a_1 \to \cdots \to a_n = x_0 ,$$

where $n > 0$ and all a_ν are in $\mathfrak{R}_o$. There is a set $A_o \subset f(a_o)$ such that

$$|f(a_o) - A_o| = |g(a_o)|.$$

Then

$$|f(a_1) + A_o| > |g(a_1)|.$$

For otherwise we would have

$$|g(a_1)| \leqq |f(a_1)| \leqq |f(a_1) + A_o| \leqq |g(a_1)| < \infty,$$

$$|A_o| = 0,$$

$$|g(a_o)| = |f(a_o) - A_o| = |f(a_o)|$$

which, in fact, is not true. Put

$$T^{a_o, a_1, A_o} f = f_1.$$

This is a T'-transformation. f_1 satisfies

$$|f_1(a_o)| = |g(a_o)|,$$

$$|f_1(a_1)| = |f(a_1) + A_o| > |g(a_1)|,$$

$$f_1(x) = f(x) \qquad (x \neq a_o, a_1).$$

Now treat f_1, a_1 in the same way as f, a_o were treated, etc. until, after n such T'-transformations, we obtain a function f_n satisfying

$$|f_n(x)| \begin{cases} = |g(x)| & (x = a_o, a_1, \cdots, a_{n-1}) \\ > |g(x)| & (x = a_n), \end{cases}$$

$$f_n(x) = f(x) \qquad (x \neq a_o, \cdots, a_n).$$

Put

$$O_2 f = f_n.$$

The number of points x for which

$$|f_n(x)| = |g(x)|$$

is greater than or equal to the number for which

$$|f(x)| = |g(x)|.$$

For every x of the second kind is also of the first kind, and

$$|f_n(a_o)| = |g(a_o)|; \qquad |f(a_o)| > |g(a_o)|.$$

Clearly then, in order that the number of points of the second kind be equal to the number of the first kind, it is necessary that $|f(a_n)| = |g(a_n)|$, and this can be true only prior to the first application of O_2.

Hence O_2 can only be applied to f a bounded number of times, and after the

last possible application a function $\tilde{f}$ is obtained which has the required properties.

5.6 We now come to the proof of Theorem 1. We assume that (2), (4) and (6) hold, and we have to show the existence of a T'-transform $\tilde{f}$ of f which satisfies (5).

CASE 1. Suppose that $\rho(\Gamma) = 0$. Then we define an operator O_3 as follows. O_3 is applicable to every f which satisfies (6) and, moreover,

$$|f(x_o)| < |g(x_o)|$$

for at least one point x_o. We may assume that O_3 is applicable to the given function f. Choose one such point x_o. The set

$$\mathfrak{R}_o = \mathfrak{R}(x_o)$$

is a section of Γ. For if $x \to y \prec \mathfrak{R}_o$ then, in view of $\rho(\Gamma) = 0$, we have

$$y \equiv x; \qquad x \prec \mathfrak{R}_o .$$

(6) implies (29). Hence there exists, by Lemma 6, a T'-transform f' of f which satisfies

$$|f'(x)| \begin{cases} \geqq |g(x)| & (x \prec \mathfrak{R}_o) \\ = |f(x)| & (x \not\prec \mathfrak{R}_o). \end{cases}$$

Put

$$O_3 f = f'.$$

Since the number of points x satisfying

$$|f'(x)| \geqq |g(x)|$$

exceeds the number of points x satisfying

$$|f(x)| \geqq |g(x)|$$

it follows that O_3 can only be applied a bounded number of times to the given function f, and after the last possible application we obtain a function $\tilde{f}$ which satisfies (5).

CASE 2. Suppose that $\rho(\Gamma) > 0$. We use induction with respect to $\rho(\Gamma)$. Choose a regular edge $a \to b$ for which $p(a)$ reaches its smallest possible value in Γ. Then

$$\mathfrak{R}_o = \mathfrak{R}(a)$$

is a section. For if there would be any points x, y satisfying

$$x \to y \prec \mathfrak{R}_o ; \qquad x \not\prec \mathfrak{R}_o$$

then

$$x \not\equiv y; \qquad p(x) < p(y) = p(a)$$

which is impossible. Once more, because of (6), (29) is true, and therefore there exists a T'-transform f' of f satisfying

$$|f'(x)| \geqq |g(x)| \qquad (x \prec \mathfrak{K}_o),$$

$$f'(x) = f(x) \qquad (x \not\prec \mathfrak{K}_o). \tag{32}$$

Now Lemma 7 shows the existence of a T'-transform f'' of f' which satisfies

$$|f''(x)| = |g(x)| \qquad (x \prec \mathfrak{K}_o - \{a\}), \tag{33}$$

$$f''(x) = f'(x) \qquad (x \not\prec \mathfrak{K}_o). \tag{34}$$

Then

$$|f''(a)| \geqq |g(a)|. \tag{35}$$

For let us suppose that

$$|f''(a)| < |g(a)| < \infty. \tag{36}$$

Then Lemma 2, applied to (33), (36), would lead to

$$|(f'', \mathfrak{K}_o)| < |(g, \mathfrak{K}_o)| \tag{37}$$

while, on the other hand, by (32) and (34),

$$|(f'', \mathfrak{K}_o)| = |(f'', \mathfrak{A}) - (f'', \mathfrak{A} - \mathfrak{K}_o)| = |(f, \mathfrak{A}) - (f, \mathfrak{A} - \mathfrak{K}_o)| = |(f, \mathfrak{K}_o)|.$$

This, together with (29), contradicts (37). Therefore (35). Any section $\mathfrak{A}$ which contains at least one point of $\mathfrak{K}_o$ contains the whole set $\mathfrak{K}_o$. For if

$$x \prec \mathfrak{A}\mathfrak{K}_o; \qquad y \prec \mathfrak{K}_o$$

then

$$y \longrightarrow x \prec \mathfrak{A}; \qquad y \prec \mathfrak{A}.$$

Therefore, for every $\mathfrak{A}$,

$$(f'', \mathfrak{A}) = (f, \mathfrak{A}),$$

$$|(f'', \mathfrak{A})| = |(f, \mathfrak{A})| \geqq |(g, \mathfrak{A})|. \tag{37a}$$

By (35) and (C), we can choose a set $A \subset f''(a)$ such that

$$|f''(a) - A| = |g(a)|. \tag{38}$$

Denote by

$$\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_m$$

all sections which have the property that

$$a \prec \mathfrak{A}_\mu; \qquad b \not\prec \mathfrak{A}_\mu.$$

$\mathfrak{K}_o$ is one of the sections $\mathfrak{A}_\mu$. For if b were in $\mathfrak{K}_o$ then we would have $a \equiv b$ which contradicts the fact that $a \to b$ is a regular edge. Let

$$\mathfrak{A}_1 = \mathfrak{K}_o .$$

Put

$$(f'', \mathfrak{A}_\mu) = A_\mu ; \qquad (g, \mathfrak{A}_\mu) = B_\mu \qquad (1 \leqq \mu \leqq m).$$

Then, by (33) and (38),

$$| A_1 = A | = | (f'', \mathfrak{K}_o - \{a\}) + (f''(a) - A) | = | (g, \mathfrak{K}_o - \{a\}) + g(a) |$$
$$= | (g, \mathfrak{K}_o) | = | B_1 |.$$

Also, from (37a),

$$| (A_\mu - A) + A | = | A_\mu | \geqq | B_\mu | \qquad (1 \leqq \mu \leqq m).$$

Hence, by Lemma 3, with $A_\mu - A$ instead of A_μ, there exists a set $A' \subset A$ and an index $\mu_o (1 \leqq \mu_o \leqq m)$ such that

$$| (A_\mu - A) + A' | \begin{cases} \geqq | B_\mu | \ (1 \leqq \mu \leqq m) & (39) \\ = | B_\mu | \ (\mu = \mu_o). & (40) \end{cases}$$

Then

$$T^{a,b,A-A'} f'' = f'''$$

is a T'-transformation. For

$$| f''(a) - (A - A') | \geqq | f''(a) - A | = | g(a) |.$$

Also, by (39),

$$| (f''', \mathfrak{A}_\mu) | \geqq | (g, \mathfrak{A}_\mu) | \qquad (1 \leqq \mu \leqq m).$$

If $\mathfrak{A}$ is a section which does not occur among the sections $\mathfrak{A}_\mu$ then

$$\textit{either} \quad a, b \prec \mathfrak{A}$$
$$\textit{or} \quad a, b \not\prec \mathfrak{A}.$$

In both cases

$$(f''', \mathfrak{A}) = (f'', \mathfrak{A}).$$

Hence for every section $\mathfrak{A}$

$$| (f''', \mathfrak{A}) | \geqq | (g, \mathfrak{A}) |.$$

In particular, by (40),

$$| (f''', \mathfrak{A}_{\mu_o}) | = | A_{\mu_o} - (A - A') | = | B_{\mu_o} | = | (g, \mathfrak{A}_{\mu_o}) |.$$

Therefore Lemma 5 is applicable, with

$$\mathfrak{A}_o = \mathfrak{A}_{\mu_o}; \qquad a_o = a; \qquad b_o = b.$$

Using the notation of Lemma 5 we conclude that

$$|(f''', \mathfrak{A}')| \geqq |(g, \mathfrak{A}')|$$

for every section $\mathfrak{A}'$ of Γ'. Now

$$\rho(\Gamma') = \rho(\Gamma) - 1.$$

For if $x \to y$ is a singular edge of Γ then there exists a closed cycle of edges of Γ:

$$x = x_o \to y = x_1 \to x_2 \to \cdots \to x_n = x. \tag{41}$$

It follows that every edge $x_{\nu-1} \to x_\nu (1 \leqq \nu \leqq n)$ is a singular edge of Γ, and as $a \to b$ is a regular edge of Γ we have

$$(x_{\nu-1}, x_\nu) \neq (a, b) \qquad (1 \leqq \nu \leqq n).$$

Hence, by definition of Γ', the relations (41) hold in Γ' as well as in Γ, and consequently $x \to y$ is a singular edge of Γ'.

If, on the other hand, $x \to y$ is a singular edge of Γ' then some cycle of the form (41) exists in Γ' and hence, a fortiori, in Γ. In other words, Γ and Γ' have the same system of singular edges. Therefore, remembering that $a \to b$ is true in Γ but not true in Γ', we find $\rho(\Gamma') = \rho(\Gamma) - 1$.

Now, according to our induction hypothesis, there is a T'-transform $\bar{f}$ of f''' which satisfies (5). $\bar{f}$ is a T'-transform with respect to Γ' and therefore, a fortiori, with respect to Γ. This completes the proof of Theorem 1 and, as we have seen, also of Theorem 2.

6. Proof of theorems 3 and 4

6.1 In this section we suppose that (2) and (4) hold.

LEMMA 8. *Let $x' \equiv x''$. Then there are T''-transforms $f^{(1)}$, $f^{(2)}$, of f which satisfy $|f^{(1)}(x'')| \leqq |f(x')| \leqq |f^{(2)}(x'')|$,*

$$f^{(1)}(x) = f^{(2)}(x) = f(x) \quad (x \not\equiv x'),$$

$$|f^{(\lambda)}(x)| = |f(\phi^{(\lambda)}(x))| \qquad (1 \leqq \lambda \leqq 2; \qquad x \equiv x')$$

where, for each λ, $\phi^{(\lambda)}(x)$ effects a permutation of the points x of $\mathfrak{N}(x')$.

PROOF. 1) Since $x'' \to \to x'$ there is a chain

$$x'' = x_o \to x_1 \to \cdots \to x_n = x'$$

where $n \geqq 0$. Then

$$x_\nu \equiv x' \quad (0 \leqq \nu \leqq n).$$

Let m be the smallest index which satisfies

$$0 \leqq m \leqq n; \qquad |f(x_m)| \leqq |f(x')|.$$

If $m = 0$ then $|f(x'')| \leqq |f(x')|$, and hence we may put $f^{(1)}(x) = f(x)$. Now suppose that $m > 0$. Then

$$|f(x_\nu)| > |f(x')| \geqq |f(x_m)| \qquad (0 \leqq \nu < m).$$

By (C) there is a set $A_{m-1} \subset f(x_{m-1})$ such that

$$|f(x_{m-1}) - A_{m-1}| = |f(x_m)|,$$

$$|f(x_m) + A_{m-1}| = |(f(x_{m-1}) - A_{m-1}) + A_{m-1}|.$$

Therefore

$$T^{x_{m-1}, x_m, A_{m-1}} f = f_1$$

is a T'''-transformation, and

$$|f_1(x_{m-1})| = |f(x_m)| \leqq |f(x')|.$$

Now treat x_{m-1}, f_1 in the same way as x_m, f, etc., until, after m such T'''-transformations, f is transformed into a function f_m which satisfies

$$|f_m(x_\nu)| = \begin{cases} |f(x_m)| & (\nu = 0) \\ |f(x_{\nu-1})| & (0 < \nu \leqq m), \end{cases}$$

$$f_m(x) = f(x) \qquad (x \neq x_0, x_1, \cdots, x_m).$$

Hence we may put

$$f^{(1)}(x) = f_m(x).$$

2) Since $x' \to\to x''$ there is a chain

$$x' = y_0 \to y_1 \to \cdots \to y_k = x''$$

where $k \geqq 0$. Let l be the largest index satisfying

$$0 \leqq l \leqq k; \qquad |f(y_l)| \geqq |f(x')|.$$

If $l = k$ then

$$|f(x'')| \geqq |f(x')|,$$

and we may put $f^{(2)}(x) = f(x)$. Now suppose that $l < k$. Then

$$|f(y_\nu)| < |f(x')| \leqq |f(y_l)| \qquad (l < \nu \leqq k).$$

Hence there is a set $B_l \subset f(y_l)$ such that

$$|f(y_l) - B_l| = |f(y_{l+1})|,$$

$$|f(y_{l+1}) + B_l| = |f(y_l)|,$$

and

$$T^{y_l, y_{l+1}, B_l} f = f_1'$$

is a T'''-transformation,

$$|f_1'(y_{l+1})| = |f(y_l)| \geqq |f(x')|.$$

Now treat y_{l+1}, f_1' in the same way as y_l, f, etc., until, after $k - l$ such T''-transformations, f is transformed into a function f_{k-l}' satisfying

$$|f_{k-l}'(y_\nu)| = \begin{cases} |f(y_{\nu+1})| & (l \leqq \nu < k) \\ |f(y_l)| & (\nu = k), \end{cases}$$

$$f_{k-l}'(x) = f(x) \qquad (x \neq y_l, \quad y_{l+1}, \cdots, y_k).$$

Thus we may put

$$f^{(2)}(x) = f_{k-l}'(x).$$

6.2 LEMMA 9. *Suppose that $g(x)$ is decreasing, $\mathfrak{K}_o = \mathfrak{K}(x_o)$,*

$$|(f, \mathfrak{K}_o)| \geqq |(g, \mathfrak{K}_o)|. \tag{42}$$

Then there exists a T''-transform $\tilde{f}$ of f satisfying

$$|\tilde{f}(x)| \geqq |g(x)| \quad (x \prec \mathfrak{K}_o)$$

$$\tilde{f}(x) = f(x) \qquad (x \not\prec \mathfrak{K}_o).$$

PROOF. If $x \prec \mathfrak{K}_o$ then $x \to\to x_o \to\to x$ and hence

$$|g(x)| \geqq |g(x_o)| \geqq |g(x)|,$$

$$|g(x)| = |g(x_o)| \qquad (x \prec \mathfrak{K}_o). \tag{43}$$

Define an operator O_4 as follows. O_4 is applicable to every f which satisfies (42) and

$$|f(x')| < |g(x')| \tag{44}$$

for at least one $x' \prec \mathfrak{K}_o$. We may assume that O_4 is applicable to the given function f. Choose one such x'. Then there exists $x'' \prec \mathfrak{K}_o - \{x'\}$ such that

$$|f(x'')| > |g(x_o)|. \tag{45}$$

For if

$$|f(x)| \leqq |g(x_o)| \qquad (x \prec \mathfrak{K}_o - \{x'\}) \tag{46}$$

were true then, by Lemma 2 and (43), (44), (46) we obtain a contradiction against (42). Choose a point $x'' \prec \mathfrak{K}_o - \{x'\}$ which satisfies (45). Then there is a chain

$$x'' = a_o \to a_1 \to \cdots \to a_n = x',$$

where $n > 0$ and all a_ν belong to $\mathfrak{K}_o$. Let ν_o be the largest index for which

$$0 \leqq \nu_o < n; \qquad |f(a_{\nu_o})| > |g(x_o)|.$$

Let ν_1 be the smallest index for which

$$\nu_o < \nu_1 \leqq n; \qquad |f(a_{\nu_1})| \neq |g(x_o)|.$$

Then

$$|f(a_\nu)| \begin{cases} = |g(x_o)| & (\nu_o < \nu < \nu_1) \\ < |g(x_o)| & (\nu = \nu_1). \end{cases}$$

In case we should have $\nu_o + 1 < \nu_1$ choose a set $A_o \subset f(a_{\nu_o})$ satisfying

$$|f(a_{\nu_o}) - A_o| = |g(x_o)|.$$

Then, by (A),

$$|f(a_{\nu_o+1}) + A_o| = |f(a_{\nu_o})|.$$

Hence

$$T^{a_{\nu_o}, a_{\nu_o+1}, A_o} f = f_1$$

is a T''-transformation. We have

$$|f_1(a_\nu)| = \begin{cases} |g(x_o)| & (\nu = \nu_o) \\ |f(a_{\nu_o})| & (\nu = \nu_o + 1) \\ f_1(x) = f(x) & (x \neq a_{\nu_o}, a_{\nu_o+1}). \end{cases}$$

In case $\nu_o + 2 < \nu_1$, apply to f_1, $\nu_o + 1$ the same process as to f, ν_o, etc., until, after $r = \nu_1 - \nu_o + 1$ such transformations, we obtain a T''-transform f_r of f which satisfies

$$|f_r(a_\nu)| = \begin{cases} |f(a_{\nu_o})| & (\nu = \nu_1 - 1) \\ |g(x_o)| & (\nu_o \leqq \nu < \nu_1 - 1), \end{cases}$$

$$f_r(x) = f(x) \qquad (x \neq a_{\nu_o}, a_{\nu_o+1}, \cdots, a_{\nu_1-1}).$$

We put $f_o = f$, and then the result holds also for $r = 0$. Let

$$a_{\nu_1-1} = a; \qquad a_{\nu_1} = b; \qquad f_r = f'.$$

Then

$$a, b < \Re_o; \qquad a \to b;$$

$$|f'(a)| > |g(x_o)|; \qquad |f'(b)| < |g(x_o)|.$$

By (C) there is a set $A \subset f'(a)$ such that

$$|f'(a) - A| = |g(x_o)|.$$

Case 1. Suppose that $|f'(b) + A| \leqq |f'(a)|$. Then

$$T^{a,b,A} f' = f''$$

is a T''-transformation, and $|f''(a)| = |g(x_o)|$. In this case put, by definition, $0_4 f = f''$.

CASE 2. Suppose that $|f'(b) + A| > |f'(a)|$. Then

$$|f'(b)| < |g(x_o)| < |f'(a)| < |f'(b) + A|.$$

Hence an application of Lemma 3, with

$$m = 1; \qquad A_1 = f'(b); \qquad B_1 = g(x_o)$$

shows the existence of a set $A' \subset A$ satisfying

$$|f'(b) + A'| = |g(x_o)|.$$

Then

$$|f'(b) + A'| < |f'(a)|,$$

and therefore

$$T^{a,b,A'}f' = f''$$

is a T''-transformation. We have

$$|f''(b)| = |g(x_o)|.$$

Again define $O_4 f$ by putting

$$O_4 f = f''.$$

In either case the number of points x for which

$$|f''(x)| = |g(x_o)|$$

exceeds the number of points x for which

$$|f(x)| = |g(x_o)|.$$

Therefore O_4 can only be applied to f a bounded number of times, and after the last possible application we obtain a function $\tilde{f}$ of the required type.

6.3 We now prove Theorem 3. Suppose that g is decreasing and that f satisfies (6), in addition to (2) and (4). We have to show the existence of a T''-transform $\tilde{f}$ of f for which (5) is true.

CASE 1. Suppose that

$$(47) \qquad \begin{cases} |f(a)| \leqq |f(b)| \\ \textit{whenever} \quad a \longrightarrow\!\!\!\!\rightarrow b \not\equiv a. \end{cases}$$

(47) is, for instance, satisfied if $\rho(\Gamma) = O$.

We define an operator O_5 as follows. O_5 is applicable to every f satisfying (6) and (47) which is, moreover, such that

$$(48) \qquad |f(x_o)| < |g(x_o)|$$

for at least one point x_o. We may assume that O_5 is applicable to the given function f. Choose one such x_o. Then

$$\mathfrak{K}_o = \mathfrak{K}(x_o)$$

is a section. For otherwise there would exist points x_1, x_2 such that

$$x_1 \rightarrow x_2 \equiv x_o \not\equiv x_1 .$$

Then

$$\mathfrak{A}_o = [\{x_o\}] - \mathfrak{K}_o$$

is a non-empty section. For, first of all, $x_1 \prec \mathfrak{K}_o$. Also, if

$$x \rightarrow y \prec \mathfrak{A}_o$$

then

$$x \rightarrow y \rightarrow\rightarrow x_o \not\equiv y, \qquad x \rightarrow\rightarrow x_o . \tag{49}$$

But $x \equiv x_o$ is impossible. For this would imply

$$x \rightarrow y \rightarrow\rightarrow x_o \rightarrow\rightarrow x$$

which would already contradict (49). Hence

$$x \rightarrow\rightarrow x_o \not\equiv x, \tag{50}$$

i.e., $x \prec \mathfrak{A}_o$. Therefore $\mathfrak{A}_o$ is a section.

Every point x of $\mathfrak{A}_o$ satisfies (50). Therefore, by (47), (48) and the fact that g is decreasing,

$$|f(x)| \leqq |f(x_o)| < |g(x_o)| \leqq |g(x)| \qquad (x \prec \mathfrak{A}_o).$$

Now Lemma 2 shows that

$$|(f, \mathfrak{A}_o)| < |(g, \mathfrak{A}_o)|$$

which contradicts (6). This contradiction was deduced from the assumption that $\mathfrak{K}_o$ was no section. Therefore $\mathfrak{K}_o$ is, in fact, a section and, in consequence, satisfies (42).

By Lemma 9 we can find a T'''-transform f' of f which is such that

$$|f'(x)| \geqq |g(x)| \qquad (x \prec \mathfrak{K}_o)$$
$$f'(x) = f(x) \qquad (x \not\prec \mathfrak{K}_o).$$

Define O_5 by putting

$$O_5 f = f'.$$

There are more points x for which

$$|f'(x)| \geqq |g(x)|$$

than points x for which

$$|f(x)| \geqq |g(x)|.$$

Hence O_5 can only be applied to f a bounded number of times, and after the last possible application we obtain a function $\bar{f}$ with the required properties.

CASE 2. Suppose that there are points a, b for which

$$a \to\to b \not\equiv a, \tag{51}$$

$$|f(a)| > |f(b)|. \tag{52}$$

In this case the proof is based on the following principle. By means of a certain process the problem presented by f, Γ will be reduced to an analogous problem f^*, Γ^*, and we shall have either

$$\text{(i)} \qquad \Gamma^* = \Gamma; \qquad s(f^*, \Gamma^*) > s(f, \Gamma)$$

or

$$\text{(ii)} \qquad \rho(\Gamma^*) < \rho(\Gamma).$$

If such reduction process has been established then the given problem can be reduced to one falling under case one, i.e. where (47) holds. For if this were not so the reduction process could be applied infinitely often and would lead to a sequence of pairs $f^{(n)}$, $\Gamma^{(n)}$ which is such that, for every $n = 1, 2, \cdots$, either

$$\text{(i)} \qquad \Gamma^{(n+1)} = \Gamma^{(n)}; \qquad s(f^{(n+1)}, \Gamma^{(n+1)} > s(f^{(n)}, \Gamma^{(n)})$$

or

$$\text{(ii)} \qquad \rho(\Gamma^{(n+1)}) < \rho(\Gamma^{(n)}).$$

In particular, $\rho(\Gamma^{(n+1)} \leqq \rho(\Gamma^{(n)})$ for all n. Since $s(f, \Gamma)$ does not exceed a constant $C(\Gamma)$ independent of f—e.g. $C(\Gamma) = N^3$ where N is the number of points of Γ—it follows that only a finite number of consecutive integers n can satisfy (i). Therefore $\rho(\Gamma^{(n)}) \to -\infty$ as $n \to \infty$, which is absurd.

In fact, in case (ii) $\Gamma^{(n+1)}$ is obtained from $\Gamma^{(n)}$ by removing one edge, and hence it follows that the number of possible applications of the reduction process is less than a constant dependent only on Γ.

In order to define the reduction process we start by noting that, according to (51), there exists a chain

$$a = a_0 \to a_1 \to \cdots \to a_m = b.$$

Then $m > 0$ since otherwise $a = b$ which contradicts (51). Let

$$\mu_1, \mu_2, \cdots, \mu_k$$

be those indices μ for which

$$1 \leqq \mu \leqq m; \qquad a_{\mu-1} \not\equiv a_\mu,$$

and let

$$1 \leqq \mu_1 < \mu_2 < \cdots < \mu_k \leqq m.$$

Then $k > 0$ since otherwise (51) would again be violated. Put

$$a_{\mu_\kappa - 1} = x_\kappa ; \qquad a_{\mu_\kappa} = y_\kappa \qquad (1 \leqq \kappa \leqq k),$$
$$y_0 = a; \qquad x_{k+1} = b.$$

Then

$$y_0 \equiv x_1 \to y_1 \equiv x_2 \to y_2 \equiv \cdots \equiv x_k \to y_k \equiv x_{k+1} ,$$
$$x_\kappa \not\equiv y_\kappa \qquad (1 \leqq \kappa \leqq k).$$

Not all inequalities

$$|f(y_0)| \leqq |f(y_1)| \leqq |f(y_2)| \leqq \cdots \leqq |f(y_{k-1})| \leqq |f(x_{k+1})|$$

are true. For this would contradict (52). Hence either there is a number κ_0 satisfying

$$1 \leqq \kappa_0 < k; \qquad |f(y_{\kappa_0 - 1})| > |f(y_{\kappa_0})| \tag{53}$$

or we have

$$|f(y_{k-1})| > |f(b)| . \tag{54}$$

In case (53) we put

$$a' = y_{\kappa_0 - 1} ; \qquad a'' = x_{\kappa_0} ; \qquad b' = b'' = y_{\kappa_0} .$$

In case (54) we put

$$a' = y_{k-1} ; \qquad a'' = x_k ; \qquad b' = x_{k+1} ; \qquad b'' = y_k .$$

Then in either case

$$a' \equiv a'' \to b'' \equiv b'; \qquad a'' \not\equiv b'',$$
$$|f(a')| > |f(b')| .$$

Apply to $f(x)$ Lemma 8, with $x' = a'$; $x'' = a''$. We obtain a T''-transform f_1 of f which satisfies

$$|f_1(a'')| \geqq |f(a')| ,$$
$$|f_1(x)| = |f(\phi^{(2)}(x))| \quad \textit{for} \quad x \equiv a' \tag{55}$$
$$f_1(x) = f(x) \quad \textit{for} \quad x \not\equiv a' \tag{56}$$

where $\phi^{(2)}(x)$ is a permutation of the points of $\mathfrak{R}(a')$. Now apply Lemma 8 to $f_1(x)$, with $x' = b'$; $x'' = b''$. We obtain a T''-transform f_2 of f_1 which satisfies

$$|f_2(b'')| \leqq |f_1(b')| ,$$
$$|f_2(x)| = |f_1(\phi^{(1)}(x))| \quad \textit{for} \quad x \equiv b',$$
$$f_2(x) = f_1(x) \quad \textit{for} \quad x \not\equiv b'$$

where $\phi^{(1)}(x)$ is a permutation of the points of $\mathfrak{R}(b')$.

Let $\mathfrak{A}$ be any section. If $a' \prec \mathfrak{A}$ then $\mathfrak{R}(a') \subset \mathfrak{A}$ and hence, by (56),

$$(f_1, \mathfrak{A}) = (f_1, \mathfrak{G}) - (f_1, \mathfrak{G} - \mathfrak{A}) = (f, \mathfrak{G}) - (f, \mathfrak{G} - \mathfrak{A}) = (f, \mathfrak{A}).$$

On the other hand, if $a' \not\prec \mathfrak{A}$ then $\mathfrak{A}\mathfrak{R}(a') = 0$ and, by (56),

$$(f_1, \mathfrak{A}) = (f, \mathfrak{A}). \tag{57}$$

Thus (57) holds for every $\mathfrak{A}$. In the same way it follows that

$$(f_2, \mathfrak{A}) = (f_1, \mathfrak{A}) \quad \textit{for all } \mathfrak{A}.$$

Therefore

$$|(f_2, \mathfrak{A})| = |(f, \mathfrak{A})| \geqq |(g, \mathfrak{A})| \tag{58}$$

for all $\mathfrak{A}$. Furthermore,

$$|f_2(x)| = |f(\phi(x))| \quad \textit{for all } x$$

where

$$\phi(x) = \begin{cases} \phi^{(2)}(x) & (x \equiv a') \\ \phi^{(1)}(x) & (x \equiv b') \\ x & (x \not\equiv a', b'). \end{cases}$$

On making use of the various properties of our points and functions we see that

$$|f_2(a'')| = |f_1(a'')| \geqq |f(a')| > |f(b')| = |f_1(b')| \geqq |f_2(b'')|. \tag{59}$$

Therefore, by (C), we can choose a set $A \subset f_2(a'')$ such that

$$|f_2(a'') - A| = |f_2(b'')|.$$

Then, by (A),

$$|f_2(b'') + A| = |f_2(a'')|.$$

Hence

$$T^{a'', b'', A} f_2 = f_3$$

is a T'''-transformation.

Case (i). • Assume that

$$|(f_3, \mathfrak{A})| \geqq |(g, \mathfrak{A})| \qquad \textit{for all } \mathfrak{A}.$$

We have

$$|f_3(a'')| = |f_2(b'')|; \qquad |f_3(b'')| = |f_2(a'')|,$$
$$f_3(x) = f_2(x) \qquad (x \neq a'', b'').$$

For the sake of simplicity we write $p(\bar{x})$, $q(f, \bar{x})$, $s(f)$ instead of $p(\bar{x}, \Gamma)$, $q(f, \bar{x}, \Gamma)$, $s(f, \Gamma)$ respectively.

From (56), (55) and the definition of $q(f, \bar{x})$ it is obvious that

$$q(f_1, \bar{x}) = q(f, \bar{x}) \qquad (\bar{x} \not\equiv a').$$

Also,

$$p(x) = p(y) \qquad \textit{for } x \equiv y,$$
$$q(f_1, \bar{x}) = q(f, \phi^{(2)}(\bar{x})) \qquad \textit{for all } \bar{x} \equiv a'.$$

Therefore

$$s(f_1) = \sum_{\bar{x} \not\equiv a'} p(\bar{x})q(f_1, \bar{x}) + \sum_{\bar{x} \equiv a'} p(\bar{x})q(f_1, \bar{x})$$
$$= \sum_{\bar{x} \not\equiv a'} p(\bar{x})q(f, \bar{x}) + \sum_{\bar{x} \equiv a'} p(\phi^{(2)}(\bar{x}))q(f, \phi^{(2)}(\bar{x})) = s(f).$$

Similarly, using b', $\phi^{(1)}$ instead of a', $\phi^{(2)}$, we see that

$$s(f_2) = s(f_1).$$

Thus

$$s(f_2) = s(f). \tag{60}$$

Furthermore,

$$q(f_3, \bar{x}) = q(f_2, \bar{x}) \qquad \text{for } \bar{x} \neq a'', b'',$$
$$p(a'') < p(b'')$$

and, by (59),

$$q(f_3, a'') = q(f_2, b'') < q(f_2, a'') = q(f_3, b'').$$

Hence

$$s(f_3) - s(f_2)$$
$$= p(a'')q(f_3, a'') + p(b'')q(f_3, b'') - p(a'')q(f_2, a'') - p(b'')q(f_2, b'')$$
$$= (p(b'') - p(a''))(q(f_2, a'') - q(f_2, b'')) > 0,$$

and finally, by (60),

$$s(f_3, \Gamma) > s(f, \Gamma).$$

This is case (i) of the reduction process mentioned on p. 34, with

$$f^* = f_3.$$

CASE (ii). Suppose there exists a section $\mathfrak{A}_1$ for which

$$|(f_3, \mathfrak{A}_1)| < |(g, \mathfrak{A}_1)|.$$

Then

$$a'' \prec \mathfrak{A}_1; \qquad b'' \not\prec \mathfrak{A}_1.$$

As in the proof of Theorem 1, let

$$\mathfrak{A}_1, \mathfrak{A}_2, \cdots, \mathfrak{A}_m$$

be all sections for which

$$a'' \prec \mathfrak{A}_\mu; \qquad b'' \not\prec \mathfrak{A}_\mu.$$

Put

$$(f_2, \mathfrak{A}_\mu) = A_\mu; \qquad (g, \mathfrak{A}_\mu) = B_\mu \qquad (1 \leqq \mu \leqq m).$$

Then

$$|A_1 - A| = |(f_2, \mathfrak{A}_1) - A| = |(f_3, \mathfrak{A}_1)| < |(g, \mathfrak{A}_1)| = B_1,$$
$$|(A_\mu - A) + A| = |A_\mu| \geqq |B_\mu| \qquad (1 \leqq \mu \leqq m) \quad \text{(by (58))}.$$

Therefore, by Lemma 3, there is a set $A' \subset A$ and an index μ_o $(1 \leqq \mu_o \leqq m)$ such that

$$|(A_\mu - A) + A'| \begin{cases} \geqq |B_\mu| & (1 \leqq \mu \leqq m) \\ = |B_\mu| & (\mu = \mu_o). \end{cases}$$

Then

$$T^{a'',b'',A-A'} f_2 = f_4$$

is a T''-transformation, and we have

$$|(f_4, \mathfrak{A}_\mu)| \begin{cases} \geqq |(g, \mathfrak{A}_\mu)| & (1 \leqq \mu \leqq m) \\ = |(g, \mathfrak{A}_\mu)| & (\mu = \mu_o). \end{cases}$$

If $\mathfrak{A}$ is a section which does not occur among the $\mathfrak{A}_\mu$ then

$$\textit{either} \quad a'', b'' \prec \mathfrak{A} \quad \textit{or} \quad a'', b'' \nprec \mathfrak{A},$$

and hence in any case

$$(f_4, \mathfrak{A}) = (f_2, \mathfrak{A}).$$

Therefore

$$|(f_4, \mathfrak{A})| \geqq |(g, \mathfrak{A})| \qquad \textit{for all } \mathfrak{A},$$
$$|(f_4, \mathfrak{A}_o)| = |(g, \mathfrak{A}_o)|$$

when

$$\mathfrak{A}_o = \mathfrak{A}_{\mu_o}.$$

Now apply Lemma 5, with

$$a_o = a''; \qquad b_o = b''.$$

Then, in the notations of Lemma 5,

$$|(f_4, \mathfrak{A}')| \geqq |(g, \mathfrak{A}')|$$

for every section $\mathfrak{A}'$ of Γ', and we have

$$\rho(\Gamma') = \rho(\Gamma) - 1.$$

This is case (ii) of the reduction process, with

$$\Gamma^* = \Gamma'; \qquad f^* = f_4.$$

Theorems 3 and 4 are proved.

7. Proof of theorems 5 and 6

7.1 We now consider the case of two arbitrary functions $f(x)$, $g(x)$, not necessarily satisfying (4). But (2) is required. It will be shown that the transformation problem of Theorem 5 is identical with that of Theorem 1, if the latter is applied to the functions $F(\mathfrak{B})$, $G(\mathfrak{B})$ (defined on p. 10) and a certain graph Γ^*. The condition (14) is, in fact, identical with (6), applied to F, G, Γ^*.

7.2 It is obvious from the definition of $F(\mathfrak{B})$ that

$$F(\mathfrak{B}_1)F(\mathfrak{B}_2) = 0 \qquad \textit{if } \mathfrak{B}_1 \neq \mathfrak{B}_2. \tag{61}$$

To prove this formally we may assume that there is a point x_o such that $x_o \prec \mathfrak{B}_1$; $x_o \nprec \mathfrak{B}_2$. Then by (10)

$$F(\mathfrak{B}_1)F(\mathfrak{B}_2) \subset f(x_o)(S - f(x_o)) = 0.$$

Denote by Γ^* a graph whose "points" are all sets $\mathfrak{B} \subset \mathfrak{G}$. The edges of Γ^* will be defined later. The symbol $\rightarrow$ refers to Γ. Edges of Γ^* will be denoted by the symbol $\overset{*}{\rightarrow}$. More generally, the addition of a star * to any of our symbols indicates that this symbol refers to Γ^*.

Suppose that

$$a \rightarrow b; \qquad A \subset f(a),$$
$$T^{a,b,A}f = f_1 .$$

We want to investigate the relation between the functions $F(\mathfrak{B})$ and $F_1(\mathfrak{B})$.

We have, by the distributive law of multiplication,

$$A = Af(a) \prod_{x \prec \mathfrak{G} - \{a\}} (f(x) + (S - f(x))) = \sum_{a \prec \mathfrak{C}} AF(\mathfrak{C}).$$

Hence, in view of (61), the transformation $T^{a,b,A}$ is equivalent to the succession of transformations

$$T^{a,b,AF(\mathfrak{C})}$$

where $\mathfrak{C}$ ranges over all sets which contain the point a. It is therefore sufficient for our purpose to assume that

$$0 \neq A \subset F(\mathfrak{C})$$

where $\mathfrak{C}$ is some fixed set and

$$a \prec \mathfrak{C}.$$

Consider any set $\mathfrak{B}$ satisfying

$$a, b \nprec \mathfrak{B}.$$

Put

$$B = \prod_{x \prec \mathfrak{B}} f(x) \prod_{y \prec \mathfrak{B} + \{a,b\}} (S - f(x)). \tag{62}$$

Then

$$F_1(\mathfrak{B}) = B(S - (f(a) - A))(S - (f(b) + A))$$
$$= B((S - f(a)) + A)(S - f(b))(S - A) = F(\mathfrak{B})(S - A) = F(\mathfrak{B}). \tag{63}$$

The last equation follows from

$$AF(\mathfrak{B}) \subset F(\mathfrak{C})F(\mathfrak{B})$$

and (61). Similarly,

$$F_1(\mathfrak{B} + \{a\}) = B(f(a) - A)(S - (f(b) + A))$$
$$= Bf(a)(S - A)(S - f(b)) = F(\mathfrak{B} + \{a\})(S - A)$$

$$(64) \qquad = \begin{cases} F(\mathfrak{B} + \{a\}) - A & \text{when } \mathfrak{B} + \{a\} = \mathfrak{C} \\ F(\mathfrak{B} + \{a\}) & \text{otherwise.} \end{cases}$$

$$F_1(\mathfrak{B} + \{b\}) = B(f(b) + A)(S - (f(a) - A)) = Bf(b)(S - f(a)) + AB$$

$$(65) \qquad = F(\mathfrak{B} + \{b\}) + AB = \begin{cases} F(\mathfrak{B} + \{b\}) + A & \text{when } A \subset B \\ F(\mathfrak{B} + \{b\}) & \text{otherwise.} \end{cases}$$

$$F_1(\mathfrak{B} + \{a, b\}) = B(f(a) - A)(f(b) + A) = Bf(a)(S - A)f(b)$$

$$(66) \quad = F(\mathfrak{B} + \{a, b\})(S - A) = \begin{cases} F(\mathfrak{B} + \{a, b\}) - A & \text{when } \mathfrak{B} + \{a, b\} = \mathfrak{C} \\ F(\mathfrak{B} + \{a, b\}) & \text{otherwise.} \end{cases}$$

I assert that $A \subset B$ if, and only if,

$$(67) \qquad \mathfrak{B} + \{b\} = (\mathfrak{C} - \{a\}) + \{b\}.$$

First we notice that $A \subset B$ if, and only if, $F(\mathfrak{C}) \subset B$. Now it follows from (62) and the distributive law of multiplication that

$$B = B(f(a) + (S - f(a)))(f(b) + (S - f(b)))$$

$$= F(\mathfrak{B}) + F(\mathfrak{B} + \{a\}) + F(\mathfrak{B} + \{b\}) + F(\mathfrak{B} + \{a, b\}).$$

Hence $A \subset B$ is equivalent to

$$(68) \qquad \mathfrak{C} = \mathfrak{B} + \mathfrak{C}' \quad \text{where} \quad \mathfrak{C}' = \{a\} \text{ or } \{a, b\}$$

Since $a \prec \mathfrak{C}$, (68) is the same as

$$\mathfrak{B} \subset \mathfrak{C} - \{a\} \subset \mathfrak{B} + \{b\}$$

which, in turn, is equivalent to (67).

If we examine the equations (63)–(66), bearing in mind that $A \subset B$ is equivalent to (67), we find that they can be summarised as follows.

$$F_1(\mathfrak{B}_1) = \begin{cases} F(\mathfrak{B}_1) - A & \text{if } \mathfrak{B}_1 = \mathfrak{C} \\ F(\mathfrak{B}_1) + A & \text{if } \mathfrak{B}_1 = (\mathfrak{C} - \{a\}) + \{b\} \\ F(\mathfrak{B}_1) & \text{otherwise.} \end{cases}$$

Hence the appropriate definition of edges in Γ^*—the graph whose "points" are all sets $\mathfrak{B}$ of points of Γ—is

$$(69) \qquad \mathfrak{C} \xrightarrow{*} (\mathfrak{C} - \{a\}) + \{b\}$$

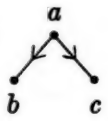
where a, b, $\mathfrak{C}$ satisfy

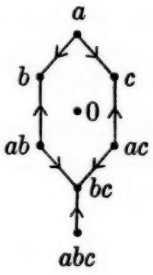
$$a \to b; \qquad a \prec \mathfrak{C}$$

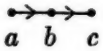
and are otherwise arbitrary. For if these are the edges of Γ^*, and we actually adopt this definition, then every T-transformation applied to $f(x)$ and the graph Γ has an effect upon the function $F(\mathfrak{B})$ which is the same as that of a succession

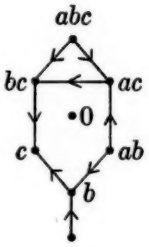
of T-transformations applied to $F(\mathfrak{B})$ and the graph Γ^*. Vice versa, our analysis shows that every T-transformation applied to $F(\mathfrak{B})$ and Γ^* is equivalent, in its effect upon $f(x)$, to a T-transformation applied to $f(x)$ and Γ.

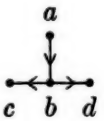
The following diagrams serve to illustrate the connection between Γ and Γ^*. Points of Γ^* which are denoted by ab, abc etc. should really be called $\{a, b\}$, $\{a, b, c\}$, etc. The point of Γ^* called 0 corresponds to the set $\mathfrak{B} = 0$. It is always an isolated point of Γ^*.

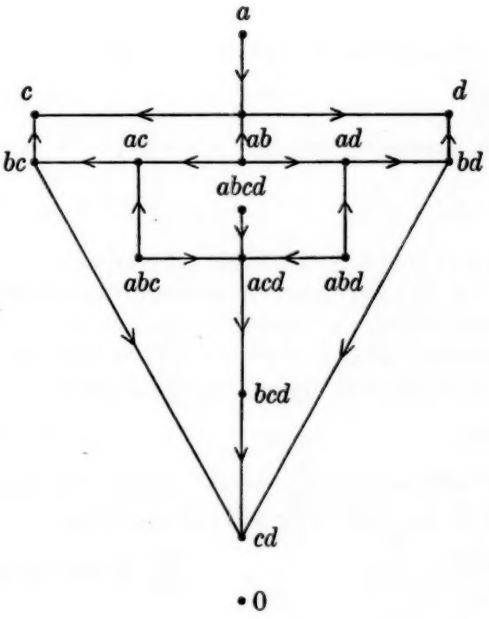
1) Γ: 

Γ^*: 

2) Γ: 

Γ^*: 

3) Γ: 

Γ^*: 

Even very simple graphs Γ give rise to quite complicated and interesting graphs Γ^*. In example 3) a few edges, e.g. $abcd \xrightarrow{*} bcd$ have been omitted in Γ^*, but the validity of relations $\mathfrak{B} \xrightarrow{*} \mathfrak{C}$ is not affected by these changes. Therefore the sections of the reproduced graph Γ^* are the same as those of the true graph Γ^*. Hence the meaning of condition (6) and thus the existence or non-existence of any T-transform mentioned in our theorems is not affected.

7.3 In order to complete the proof of theorems 5 and 6, all we have to do is to find the sections of Γ^* and to verify that (14) is the same as (6) if the latter is applied to Γ^* instead of Γ. We have to note here that

$$(f, \mathfrak{G}) = \sum_{\mathfrak{B}} F(\mathfrak{B}) = (F, \mathfrak{G}^*)^*,$$

and that therefore (8) serves for both the problem on Γ and that on Γ^*.

The definition (69) can be expressed as follows. Let x_o be an arbitrary point of Γ, and $\phi_o(x_o)$ a point such that

$$x_o \rightarrow \phi_o(x_o).$$

Put

$$\phi_o(x) = x \qquad \text{for } x \neq x_o.$$

Then, if $x_o \prec \mathfrak{B}_o$, we have

$$\mathfrak{B}_o \xrightarrow{*} \sum_{x \prec \mathfrak{B}_0} \{\phi_o(x)\}.$$

Therefore, for any $\mathfrak{B}$,

$$\mathfrak{B} \xrightarrow{*} \sum_{x \prec \mathfrak{B}} \{\phi_o(x)\}.$$

More generally, let $\phi(x)$ be defined for all x and denote a point of Γ for which

$$x \longrightarrow \phi(x) \qquad \text{for all } x. \tag{70}$$

Then every $\mathfrak{B}$ satisfies

$$\mathfrak{B} \xrightarrow{*} \sum_{x \prec \mathfrak{B}_0} \{\phi(x)\},$$

and in this way we obtain all relations $\mathfrak{B} \xrightarrow{*} \mathfrak{C}$.

7.4 We now want to determine all sections of Γ^* which can be written in the form $[\{\mathfrak{B}\}]^*$, i.e. which are "generated" by a single point $\mathfrak{B}$ of Γ^*. By (9) any arbitrary section of Γ^* is a sum of such special sections. We are given a set $\mathfrak{B}$, and we have to find all sets $\mathfrak{C}$ for which

$$\mathfrak{C} \xrightarrow{*} \mathfrak{B}. \tag{71}$$

In what follows the set $\mathfrak{B}$ is fixed. Let $\psi_{\mathfrak{B}}(x)$ be a function of the type defined on p. 11. Let $\mathfrak{C}$ be any set satisfying

$$\sum_{x \prec \mathfrak{B}} \{\psi_{\mathfrak{B}}(x)\} \subset \mathfrak{C} \subset [\mathfrak{B}]. \tag{72}$$

I am going to show now that sets $\mathfrak{C}$ which satisfy (72) for some $\psi_{\mathfrak{B}}(x)$, and no other sets, have the property (71).

1) Assume that (72) is true for a certain function $\psi_{\mathfrak{B}}(x)$ and a set $\mathfrak{C}$. Then we define a function $\phi(x)$ as follows.

$$\phi(x) = x \qquad \text{for } x \not\prec \mathfrak{C}$$

$$\phi(\psi_{\mathfrak{B}}(y)) = y \qquad \text{for } y \prec \mathfrak{B}.$$

Finally, if x belongs to the set

$$\mathfrak{C} - \sum_{y \prec \mathfrak{B}} \{\psi_{\mathfrak{B}}(y)\}$$

then, by (72),

$$x \prec [\mathfrak{B}].$$

Hence, by definition of $[\mathfrak{B}]$,

$$x \longrightarrow x'$$

for at least one x' of $\mathfrak{B}$. Choose one such x' and put

$$\phi(x) = x'.$$

It follows from the properties of $\psi_{\mathfrak{B}}$ that the conditions imposed upon ϕ do not contradict each other and that ϕ satisfies (70). Furthermore,

$$\sum_{x \prec \mathfrak{C}} \{\phi(x)\} = \mathfrak{B}.$$

Therefore (71) follows.

2) Let us now suppose that a set $\mathfrak{C}$ satisfies (71). Then, by definition of $\overset{*}{\longrightarrow}$, there exists a function $\phi(x)$ with property (70) for which

$$\sum_{x \prec \mathfrak{C}} \{\phi(x)\} = \mathfrak{B}. \tag{73}$$

Hence, given any point y of $\mathfrak{B}$, there is at least one $\bar{x}$ of $\mathfrak{C}$ satisfying $\phi(\bar{x}) = y$. Choose one such $\bar{x}$ and put

$$\psi(y) = \bar{x}.$$

Then

$$\psi(y) \prec \mathfrak{C}; \qquad \phi(\psi(y)) = y.$$

Clearly,

$$\psi(y_1) \neq \psi(y_2) \qquad \text{if } y_1 \neq y_2$$

and, by (70),

$$\psi(y) \longrightarrow y$$

when y, y_1, y_2 are in $\mathfrak{B}$. This means that $\psi(y)$ is one of those functions which may be denoted by $\psi_{\mathfrak{B}}(y)$. Also, in view of (70), (73), every x of $\mathfrak{C}$ belongs to $[\mathfrak{B}]$. Hence (72) holds.

7.5 Put

$$A = (F, [\{\mathfrak{B}\}]^*)^*.$$

From a remark made at the beginning of 7.4 it is obvious that our proof will be completed if we have established that, for every $\mathfrak{B} \neq 0$, A is equal to $F^*(\mathfrak{B})$ as defined by (12). For then (14) will be identical with (6), the latter applied to F, G, Γ^*.

We have, in fact,

$$A = \sum_{\mathfrak{C} \to^* \mathfrak{B}} F(\mathfrak{C}) = \sum_{\psi_{\mathfrak{B}}} \sum_{\mathfrak{C}_\psi} F(\mathfrak{C}_\psi), \tag{74}$$

where $\mathfrak{C}_\psi$, for every fixed $\psi_{\mathfrak{B}}$, ranges over all sets $\mathfrak{C}$ which satisfy (72). Put

$$\sum_{x \prec \mathfrak{B}} \{\psi_{\mathfrak{B}}(x)\} = \mathfrak{B}_\psi .$$

Then, for every fixed $\psi_{\mathfrak{B}}$,

$$\begin{aligned}
\sum_{\mathfrak{C}_\psi} F(\mathfrak{C}_\psi) &= \sum_{\mathfrak{C}_\psi} \Big(\prod_{x \prec \mathfrak{C}_\psi} f(x)\Big)\Big(\prod_{y \text{ not in } \mathfrak{C}_\psi} (S - f(y))\Big) \\
&= \sum_{\mathfrak{C}_\psi} \Big(\prod_{x \prec \mathfrak{B}_\psi} f(x)\Big)\Big(\prod_{x \prec \mathfrak{C}_\psi - \mathfrak{B}_\psi} f(x)\Big)\Big(\prod_{y \prec [\mathfrak{B}] - \mathfrak{C}_\psi} (S - f(y))\Big)\Big(\prod_{y \text{ not in } [\mathfrak{B}]} (S - f(y))\Big) \\
&= \Big(\prod_{x \prec \mathfrak{B}_\psi} f(x)\Big)\Big(\prod_{y \text{ not in } \mathfrak{B}} (S - f(y))\Big) \sum_{\mathfrak{B}_\psi \subset \mathfrak{C} \subset [\mathfrak{B}]} \Big(\prod_{x \prec \mathfrak{C} - \mathfrak{B}_\psi} f(x)\Big)\Big(\prod_{y \prec [\mathfrak{B}] - \mathfrak{C}} (S - f(y))\Big) \\
&= \Big(\prod_{x \prec \mathfrak{B}} f(\psi_{\mathfrak{B}}(x))\Big)\Big(\prod_{y \text{ not in } [\mathfrak{B}]} (S - f(x))\Big) S.
\end{aligned} \tag{75}$$

The sum in (75) has the value S in view of the distributive law. If we substitute in (74) we find that

$$A = \sum_{\psi_{\mathfrak{B}}} \Big(\prod_{x \prec \mathfrak{B}} f(\psi_{\mathfrak{B}}(x))\Big)\Big(\prod_{y \text{ not in } [\mathfrak{B}]} (S - f(y))\Big) = F^*(\mathfrak{B}).$$

8. Necessity and independence of the hypotheses of the theorems

8.1 Let Γ_o be the graph with $\mathfrak{G}_o = \{a, b, c\}$ whose edges are $a \to b$; $a \to c$. Assume that a measure $|A|$ is defined in S but that none of the properties (M), (A), (C), (D) are known to hold. Suppose that the following version of Theorem 2 holds.

Assumption α. Given any functions $f(x)$, $g(x)$, for $x \prec \mathfrak{G}_o$, and supposing that

$$f(x)f(y) = g(x)g(y) = 0 \qquad \text{for } x \neq y$$

Then a necessary and sufficient condition for f to have a T-transform $\bar{f}$ satisfying

$$|\bar{f}(x)| = |g(x)| \qquad \text{for all } x \tag{76}$$

is that

$$|(f, \mathfrak{A})| \geqq |(g, \mathfrak{A})| \qquad \textit{for all } \mathfrak{A} \textit{ of } \Gamma_o \tag{77}$$

and

$$|(f, \mathfrak{G}_o)| = |(g, \mathfrak{G}_o)|. \tag{78}$$

We are going to prove now that, under these circumstances, the measure has the properties (M), (A), (C), (D) and, moreover, that S contains only sets of finite measure. Thus all our assumptions concerning the measure are necessary for the truth of Theorem 2. The slight discrepancy due to the fact that here both, $f(x)$ and $g(x)$, are of finite measure while in Theorem 2 only $|g(x)| < \infty$, can be removed.

The non-empty sections of Γ_o are

$$\{a\}, \quad \{a, b\}, \quad \{a, c\}, \quad \{a, b, c\}.$$

For simplicity a symbol

$$f = \begin{pmatrix} B, & C \\ & A \end{pmatrix}$$

stands for the definition of a function $f(x)$, namely for

$$f(a) = A; \quad f(b) = B; \quad f(c) = C.$$

1) Let $A \subset B$. Put

$$f = \begin{pmatrix} 0, & 0 \\ B & \end{pmatrix}; \quad g = \begin{pmatrix} B - A, & 0 \\ A & \end{pmatrix}.$$

Then $T^{a,b,B-A}f = g$. Therefore, by α,

$$|B| = |f(a) + f(c)| \geqq |g(a) + g(c)| = |A|.$$

Hence property (M) holds.

2) Let

$$|A_1| = |B_1|; \quad |A_2| = |B_2|; \quad A_1A_2 = B_1B_2 = 0.$$

Put

$$f = \begin{pmatrix} A_1, & A_2 \\ 0 & \end{pmatrix}; \quad g = \begin{pmatrix} B_1, & B_2 \\ 0 & \end{pmatrix}.$$

Then

$$T^{a,b,0}f = f; \quad |f(x)| = |g(x)| \qquad \textit{for all } x.$$

Hence, by α,

$$|A_1 + A_2| = |(f, \mathfrak{G}_o)| = |(g, \mathfrak{G}_o)| = |B_1 + B_2|.$$

Therefore (A) holds.

3) Let

$$|A| \geqq |B_1|; \qquad A - AB_1 = A_1; \qquad B_1 - AB_1 = B_2.$$

Put

$$f = \begin{pmatrix} B_2, & 0 \\ A & \end{pmatrix}; \qquad g = \begin{pmatrix} A_1, & 0 \\ B_1 & \end{pmatrix}.$$

Then (77) and (78) are satisfied. Hence, by α, there exist sets C, D satisfying

$$CD = 0; \qquad C + D \subset A$$

which are such that (76) holds when

$$\bar{f} = \begin{pmatrix} B_2 + C, & D \\ A - (C + D) & \end{pmatrix}.$$

Then

$$|A - (C + D)| = |\bar{f}(a)| = |g(a)| = |B_1|.$$

This establishes (C).

4) Let

$$|A| \geqq |B_1 + B_2|; \qquad B_1B_2 = 0.$$

By (C) there is $A' \subset A$ such that $|A'| = |B_1 + B_2|$.
Put

$$f = \begin{pmatrix} 0, & 0 \\ A' & \end{pmatrix}; \qquad g = \begin{pmatrix} B_1, & B_2 \\ 0 & \end{pmatrix}.$$

Then, using (M), we deduce (77), (78). Hence, by α, there exist sets A_1, A_2 satisfying

$$A_1A_2 = 0; \qquad A_1 + A_2 \subset A'$$

which are such that (76) holds for

$$\bar{f} = \begin{pmatrix} A_1, & A_2 \\ A' - (A_1 + A_2) & \end{pmatrix}.$$

Then

$$|A_1| = |\bar{f}(b)| = |g(b)| = |B_1|; \qquad |A_2| = |B_2|.$$

Therefore (D) holds.

5) Let

$$A' \subset A; \qquad |A'| = |A|.$$

Put

$$f = \begin{pmatrix} A - A', & 0 \\ A' & \end{pmatrix}; \qquad g = \begin{pmatrix} 0, & 0 \\ A & \end{pmatrix}.$$

Then (77) and (78) hold. There are sets B, C with

$$BC = 0; \qquad B + C \subset A',$$

which are such that (76) is true for

$$\tilde{f} = \begin{pmatrix} (A - A') + B, & C \\ A' - (B + C) & \end{pmatrix}.$$

Therefore, by (M),

$$| A - A' | \leqq | (A - A') + B | = | \tilde{f}(b) | = | g(b) | = 0.$$

This result shows that all sets are of finite measure.

8.2 Let Γ be an arbitrary finite graph. We shall now investigate the independence of the inequalities (6) in the case when the values of $f(x)$, $g(x)$ are real non-negative numbers. Strictly speaking, we take as S the set of all real numbers, call measureable all subsets A of S which consist of a finite number of intervals $\alpha_\nu \leqq t < \beta_\nu$ (where $\alpha_\nu < \beta_\nu$), and we write $| A | < | A' |$ if, and only if,

$$\sum_\nu (\beta_\nu - \alpha_\nu) < \sum_\nu (\beta'_\nu - \alpha'_\nu).$$

It is obvious that in the case of non-overlapping sets $f(x)$ and $g(x)$ all operations with the sets correspond to analogous operations with their measures, and vice versa. Hence we may, for the rest of this paper, consider $f(x)$, $g(x)$ as functions defined for all points x of Γ and having as functional values real, non-negative numbers. Put

$$f(x) - g(x) = h(x).$$

Then (6) corresponds to

$$(h, \mathfrak{A}) \geqq 0 \qquad \textit{for all } \mathfrak{A}. \tag{79}$$

We want to study the interdependencies between the various inequalities (79).

Let

$$\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \cdots, \mathfrak{A}^{(m)}$$

be all sections of Γ,

$$\mathfrak{A}^{(\mu)} = \{x_1^{(\mu)}, x_2^{(\mu)}, \cdots, x_{n_\mu}^{(\mu)}\} \qquad (1 \leqq \mu \leqq m) \tag{80}$$

where, for every μ, the points

$$x_1^{(\mu)}, \cdots, x_{n_\mu}^{(\mu)}$$

are different from each other. The sections $\mathfrak{A}^{(\mu)}$ are also supposed to be different from each other. (79) is the same as

$$h(x_1^{(\mu)}) + h(x_2^{(\mu)}) + \cdots + h(x_{n_\mu}^{(\mu)}) \geqq 0 \qquad (1 \leqq \mu \leqq m). \tag{81}$$

Here the symbols $h(x)$ denote independent variables which can take any real values.

8.3 Define a relation

$$a \sim b$$

as having the following meaning. There exists an ordered system

$$x_0, x_1, \cdots, x_n$$

of points satisfying $x_0 = a$; $x_n = b$ which has the property that, for every $\nu (1 \leqq \nu \leqq n)$, at least one of the relations

$$x_{\nu-1} \to x_\nu ; \qquad x_\nu \to x_{\nu-1}$$

holds. In particular, $x \sim x$ for every x.

A set $\mathfrak{B}$ is called *connected* if

$$a \sim b \quad \textit{whenever} \quad a, b \prec \mathfrak{B}.$$

Obviously, every set $\mathfrak{B}$ can be represented as sum

$$\mathfrak{B}_1 + \mathfrak{B}_2 + \cdots + \mathfrak{B}_k$$

of connected sets $\mathfrak{B}_\kappa$ which have mutually no points in common, and this representation is, apart from the order, unique. If $\mathfrak{B}$ is a section then every $\mathfrak{B}_\kappa$ is a section. Also

$$(h, \mathfrak{B}) = \sum_{\kappa=1}^{k} (h, \mathfrak{B}_\kappa).$$

It is therefore evident that, in (81), we need only retain those inequalities which correspond to connected sections $\mathfrak{A}^{(\mu)}$. Our next theorem asserts that the remaining inequalities are, in fact, independent from each other. Let

$$\mathfrak{A}^{(1)}, \mathfrak{A}^{(2)}, \cdots, \mathfrak{A}^{(r)}$$

be those non-empty sections which are connected. Then $1 \leqq r \leqq m$.

THEOREM 7. *The inequalities*

$$(h, \mathfrak{A}^{(\rho)}) \geqq 0 \qquad (1 \leqq \rho \leqq r) \tag{82}$$

are independent from each other, and every inequality (79) *is a consequence of* (82).

PROOF. Put, for $1 \leqq \rho \leqq r$,

$$\chi^{(\rho)}(x) = \begin{cases} 1 & (x \prec \mathfrak{A}^{(\rho)}) \\ 0 & (x \nprec \mathfrak{A}^{(\rho)}). \end{cases}$$

Then (82) can be written in the form

$$L^{(\rho)} = \sum_x \chi^{(\rho)}(x) h(x) \geqq 0 \qquad (1 \leqq \rho \leqq r). \tag{83}$$

The numbers $h(x)$ are to be considered as independent real variables. Let us assume that the inequalities (83) are not independent from each other. Then $r > 1$. We can number the $\mathfrak{A}^{(\rho)}$ in such a way that the inequality

$$L^{(1)} \geqq 0$$

is a consequence of the system

$$L^{(\rho)} \geqq 0 \qquad (1 < \rho \leqq r).$$

According to a theorem which follows most naturally from the theory of convex sets of points (For an elementary proof see **3,** Satz 3. A very similar theorem is proved in **4**), this can only be true if $L^{(1)}$ is a linear combination of $L^{(2)}, \cdots, L^{(r)}$ with non-negative coefficients i.e. if there exists an identity in the variables $h(x)$ which is of the form

$$\sum_x \chi^{(1)}(x)h(x) = \sum_{\rho=2}^{r} K_\rho \sum_x \chi^{(\rho)}(x)h(x)$$

where

$$K_2, \cdots, K_r \geqq 0, \text{ const.}$$

Then

$$\chi^{(1)}(x) = \sum_{\sigma=2}^{r} K_\rho \chi^{(\rho)}(x) \qquad \textit{for all } x. \tag{84}$$

$\mathfrak{A}^{(1)}$ is not empty. Let $x_1 \prec \mathfrak{A}^{(1)}$. From

$$1 = \chi^{(1)}(x_1) = \sum_{\rho=2}^{r} K_\rho \chi^{(\rho)}(x_1)$$

and

$$K_\rho \chi^{(\rho)}(x_1) \geqq 0 \qquad (1 < \rho \leqq r) \tag{85}$$

follows that at least one of the numbers (85) is positive. Hence we may assume that

$$K_2 > 0; \qquad \chi^{(2)}(x_1) = 1. \tag{86}$$

If $x \prec \mathfrak{A}^{(2)}$ then, by (84),

$$\chi^{(1)}(x) \geqq K_2 \chi^{(2)}(x) > 0,$$

$$\chi^{(1)}(x) = 1; \qquad x \prec \mathfrak{A}^{(1)}.$$

Hence $\mathfrak{A}^{(2)} \subset \mathfrak{A}^{(1)}$. Because of $\mathfrak{A}^{(2)} \neq \mathfrak{A}^{(1)}$ there exists a point

$$x_2 \prec \mathfrak{A}^{(1)} - \mathfrak{A}^{(2)}.$$

Since

$$x_1 , x_2 \prec \mathfrak{A}^{(1)}$$

and $\mathfrak{A}^{(1)}$ is connected we can find a chain

$$x_1 = a_o ; \quad a_1 ; \quad \cdots ; \quad a_l = x_2 \tag{87}$$

which is such that, for every λ satisfying $1 \leqq \lambda \leqq l$, at least one of the relations

$$a_{\lambda-1} \to a_\lambda ; \qquad a_\lambda \to a_{\lambda-1}$$

is true. Furthermore, $a_o , a_1 , \cdots , a_l$ are in $\mathfrak{A}^{(1)}$. In view of

$$a_o \prec \mathfrak{A}^{(2)}; \qquad a_l \nprec \mathfrak{A}^{(2)}$$

there is at least one index λ_o satisfying

$$1 \leqq \lambda_o \leqq l; \qquad a_{\lambda_o-1} \prec \mathfrak{A}^{(2)}; \qquad a_{\lambda_o} \nprec \mathfrak{A}^{(2)}.$$

Remembering that $\mathfrak{A}^{(2)}$ is a section we conclude that $a_{\lambda_o} \to a_{\lambda_o-1}$ is impossible, and that therefore, according to the property of the system (87), $a_{\lambda_o-1} \to a_{\lambda_o}$. Then, using again the definition of sections, we find that

$$\chi^{(\rho)}(a_{\lambda_o-1}) - \chi^{(\rho)}(a_{\lambda_o}) \begin{cases} = & 0 \ (\rho = 1) \\ = & 1 \ (\rho = 2) \\ \geqq & 0 \ (2 < \rho \leqq r). \end{cases}$$

Finally, by (84),

$$0 = \chi^{(1)}(a_{\lambda_o-1}) - \chi^{(1)}(a_{\lambda_o}) = \sum_{\rho=2}^{r} K_\rho(\chi^{(\rho)}(a_{\lambda_o-1}) - \chi^{(\rho)}(a_{\lambda_o})) \geqq K_2$$

which contradicts (86).

In conclusion I should like to add that it is possible, by means of the same theorem on linear inequalities used above, to prove the special case of Theorem 1 where T' is replaced by T and the measure is that defined in 7.2.

THE UNIVERSITY
SHEFFIELD

REFERENCES

(1) R. RADO, *A theorem on general measure functions*, Proc. Lond. Math. Soc. 44 (1938) 61–91.
(2) HARDY, LITTLEWOOD AND POLYA, Inequalities (Cambridge 1934).
(3) H. WEYL, *Elementare Theorie der konvexen Polyeder*, Comment. Math. Helvetici 7 (1935), 290–306.
(4) E. STIEMKE, *Lineare Ungleichungen*, Math. Annalen 76 (1915), 340–342.

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

FAMILY OF TOTALLY UMBILICAL HYPERSURFACES IN AN EINSTEIN SPACE

BY YUNG-CHOW WONG[1]

(Received October 5, 1942; Revised January 4, 1943 and February 10, 1943)

1. Introduction

Let V_n be a Riemannian n-space with fundamental tensor[2] $g_{ij}(h, i, j, k, l = 1, \cdots, n)$.[3] As usual we denote by a comma the covariant differentiation with respect to g_{ij}, and by R^l_{ijk} and R_{ij} the Riemann and Ricci tensors for g_{ij}, respectively. The *scalar curvature*[4] a and the tensor L_{ij} are defined by

$$a = -\frac{R}{n(n-1)}, \tag{1.1}$$

$$L_{ij} = R_{ij} - \frac{R}{2(n-1)} g_{ij}, \tag{1.2}$$

where $R = g^{ij}R_{ij}$. V_n is called an *Einstein* space, a *conformal-Euclidean* space,[5] or a space of *constant curvature*, and is denoted by E_n, C_n, or S_n, respectively, if V_n satisfies the respective condition:[6]

$$E_n(n > 2) \qquad R_{ij} = -(n-1)ag_{ij}, \tag{1.3}$$

$$C_n(n > 2) \qquad \begin{cases} R^l_{ijk} = \dfrac{1}{n-2}(L_{i[j}\delta^l_{k]} + g_{i[j}L^l_{k]}) & \text{for } n > 3, \\ L_{i[j,k]} = 0 & \text{for } n = 3, \end{cases} \tag{1.4}$$

$$S_n(n \geqq 2) \qquad R^l_{ijk} = -ag_{i[j}\delta^l_{k]}. \tag{1.5}$$

In (1.3) and in (1.5) for $n > 2$ the scalar curvature a is automatically constant, while in (1.5) for $n = 2$, a is assumed to be constant. For $n = 3$, $(1.4)_1$ is

[1] This paper was written while the author was a Chinese Ying-Keng Scholarship Student visiting Massachusetts Institute of Technology. The author wishes to thank Prof. D. J. Struik for the conversations they had from time to time during the preparation of the manuscript.

[2] Fundamental tensors are always supposed to be non-singular, though not necessarily definite. All functions appearing in this paper are real and are assumed to have differentiability properties adequate to the part they play in the discussion.

[3] An index has the same range throughout this paper. An index which appears twice in an expression, once as superscript and once as subscript, is to be summed over the appropriate range. A numerical index at the upper right-hand corner means an exponential except in the case of the coordinates x^a, x^k, x^b or x^p.

[4] Here we follow Schouten's definition; Eisenhart calls R the scalar curvature.

[5] I.e. a V_n which can be mapped conformally on a Euclidean space. In this paper whenever we speak of a C_n, it is understood that $n > 2$. For the theory of a C_n, see, e.g., Eisenhart **5**, 89–92; Schouten-Struik **11**, 199–205, as listed at the end of this paper.

[6] We write, e.g., $L_{ij}\delta^l_k - L_{ik}\delta^l_j = L_{i[j}\delta^l_{k]}$.

identically satisfied; for $n > 3$, $(1.4)_2$ is a consequence of $(1.4)_1$. From (1.3), (1.4) and (1.5), it follows at once that an $S_n (n > 2)$ is an E_n and a C_n; and a V_n which is an E_n and a C_n must be an S_n.

The present paper is the result of an attempt to generalize a theorem given in a previous paper of mine (Wong **14,** Th. 4.2′). Briefly stated, the theorem is: *A one-parameter family of E_n which are conformal to one another can in general be imbedded in an E_{n+1} as totally umbilical hypersurfaces.* A hypersurface in a Riemannian space is called *totally umbilical* if its first and second fundamental tensors differ by a scalar factor; in particular, it is called *totally geodesic* if its second fundamental tensor is identically zero.

In Part I of this paper we first establish a generalization of the above theorem. Then a necessary and sufficient condition is obtained for a V_n to be imbedable in an E_{n+1} as a member of ∞^1 totally umbilical hypersurfaces. The condition is that V_n have constant scalar curvature and that there exist scalar functions ρ of a certain nature satisfying the differential equation

$$\rho_{,ij} + \phi(\rho) L_{ij} \backsim g_{ij}. \tag{1.6}$$

$\phi(\rho)$ is a known polynomial in ρ, and the sign $\backsim$ indicates that the tensor on its left is equal to the tensor on its right multiplied by a scalar. The determination of the imbedding E_{n+1}, if it exists, depends solely on the solutions of (1.6). An application of a classical theorem on a system of linear differential equation reduces the compatibility of (1.6) to that of a system of polynomial equations.

A desire to determine all the C_n imbedable in an E_{n+1} as a member of ∞^1 totally umbilical hypersurfaces leads to the work of Part II, which contains a detailed study of (1.6) for the case of a C_n. For an E_n or a V_2, equation (1.6) reduces to $\rho_{,ij} \backsim g_{ij}$, which has been studied by Brinkmann (**2**), Fialkow (**8**), Yano (**13**), and Wong (**14,** §2). Here it is proved that a C_n, not an S_n, for which equation (1.6) admits a solution is characterized by its tensor L_{ij} having one of two particular forms. Canonical forms for the fundamental tensors of such C_n are obtained.

I. Totally Umbilical Hypersurfaces in an E_{n+1}

2. Family of a totally umbilical V_n in an E_{n+1}

Given in a V_m $(m = n + 1 > 2)$ a one-parameter family of hypersurfaces V_n, whose (first) fundamental tensors are of rank n, then a coordinate system exists in which the equation of V_n is $x^m = \text{const.}$ and the fundamental tensor ${}^*g_{\alpha\beta}(\alpha, \beta = 1, \cdots, m)$ is such that ${}^*g_{im} = 0 (i = 1, \cdots, n)$ (Eisenhart **5,** 144 and 5). It is known (Eisenhart **5,** 182) that in this coordinate system, V_n are totally umbilical in V_m if and only if ${}^*g_{ij}$ is of the form ${}^*g_{ij} = \rho^{-2} g_{ij}$, where $\rho = \rho(x^\alpha)$ and $g_{ij} = g_{ij}(x^k)$, and that[7] $\partial_m \rho = 0$ or $\neq 0$ according as V_n are totally geodesic or not.

[7] We write $\partial_\alpha = \dfrac{\partial}{\partial x^\alpha}$.

Let the fundamental tensor of V_m be

$$(2.1) \qquad {}^*g_{\alpha\beta} = \begin{bmatrix} \rho^{-2} g_{ij} & 0 \\ 0 & {}^*g_{mm} \end{bmatrix}.$$

The fundamental tensors of V_n are evidently $\rho^{-2} g_{ij}$. In order that they may be isometric to a given set of fundamental tensors $[{}'\rho({}'x^\alpha)]^{-2} {}'g_{ij}({}'x^k)$ with ${}'x^m$ as parameter, it is necessary and sufficient that there exists a transformation of the form

$$x^k = x^k({}'x^i), \qquad x^m = x^m({}'x^m),$$

identifying these two sets of fundamental tensors. In the coordinates ${}'x^\alpha$, the fundamental tensor of V_m is of the form (2.1) with ρ, g_{ij}, ${}^*g_{mm}$ replaced by ${}'\rho$, ${}'g_{ij}$, ${}'^*g_{mm}$, respectively.

The above observations furnish us a way to discuss the imbedability of a one-parameter family of conformal V_n in an E_{n+1} as totally umbilical hypersurfaces. In fact, we may suppose the tensors $\rho^{-2} g_{ij}$ in (2.1) given, and determine the condition in which ${}^*g_{mm}$ may be found so that the tensor (2.1) represents[8] an E_{n+1}. The result is the following

THEOREM 2.1. *A one-parameter family of V_n with fundamental tensors ${}^*g_{ij}(x^k, x^m)$ and parameter x^m can be imbedded in an E_{n+1} as totally umbilical hypersurfaces, which are not totally geodesic, if and only if the following conditions are satisfied:*

(1) *${}^*g_{ij}$ is of the form $[\rho(x^k, x^m)]^{-2} g_{ij}(x^k)$, where $\partial_m \rho \neq 0$;*

(2) *Each V_n has constant scalar curvature;*

(3) *When it is assumed*[9] *that $\rho(x^k, x_o^m) = 1$ for some value x_o^m of x^m, the tensor*

$$\partial_m \left(\frac{\rho_{,ij}}{\rho^{n-1}} \right) + \frac{\partial_m \rho}{\rho^{n-1}} R_{ij}$$

differs from g_{ij} by a scalar factor.

If $n = 2$, the tensor in (3) can be replaced by $\rho_{,ij}$; if $n > 2$, it can be replaced by

$$\rho_{,ij} + \frac{\rho^{n-1} - \rho}{n-2} R_{ij} \qquad \text{and by} \qquad {}'R_{ij} - \rho^{n-1} R_{ij},$$

where ${}'R_{ij}$ is the Ricci tensor of ${}^*g_{ij}$.

---

[8] By this we mean that ${}^*g_{\alpha\beta}$ is the fundamental tensor of an E_{n+1}.

[9] This assumption is not a restriction; for, if x_o^m is any fixed value of x^m,

$$[\rho(x^k, x^m)]^{-2} g_{ij}(x^k) = \left[\frac{\rho(x^k, x^m)}{\rho(x^k, x_o^m)} \right]^{-2} \{ [\rho(x^k, x_o^m)]^{-2} g_{ij}(x^k) \},$$

and $\dfrac{\rho(x^k, x^m)}{\rho(x^k, x_o^m)}$, $[\rho(x^k, x_o^m)]^{-2} g_{ij}(x^k)$ can be used in place of $\rho(x^k, x^m)$, $g_{ij}(x^k)$, respectively. From this it follows that any tensor of the family $\rho^{-2} g_{ij}$ may be made to play the part of g_{ij}.

Let a one-parameter family of V_n with fundamental tensors ${}^*g_{ij}(x^k, x^m)$ satisfy conditions (1), (2) *and* (3), *and let the scalar curvatures of V_n be $a(x^m)$. Then if c is any constant not identically equal to $a(x^m)$, there exists one and only one E_{n+1} whose scalar curvature is c and in which V_n can be imbedded as totally umbilical hypersurfaces; the components of the fundamental tensor ${}^*g_{\alpha\beta}$ of the E_{n+1} are*

$$ {}^*g_{im} = 0, \qquad {}^*g_{ij} = \rho^{-2} g_{ij}, \qquad {}^*g_{mm} = \frac{(\partial_m \log \rho)^2}{a(x^m) - c}. $$

PROOF. The components of the Ricci tensor ${}^*R_{\alpha\beta}$ of the fundamental tensor (2.1) are (Wong **14**, (3.6))

$$ (2.2) \quad \begin{cases} {}^*R_{im} = -(m-2)(\partial_i \rho_m + \rho_m \sigma_i), \\ {}^*R_{ij} = {}'R_{ij} + \sigma\left(\frac{1}{\sigma}\right)_{;ij} + {}^*g_{ij}\,{}^*g^{mm}\left[\rho\left(\frac{1}{\rho}\right)_{;mm} + (m-2)\rho^2\left(\frac{1}{\rho}\right)_{;m}\left(\frac{1}{\rho}\right)_{;m}\right], \\ {}^*R_{mm} = (m-1)\rho\left(\frac{1}{\rho}\right)_{;mm} + {}^*g_{mm}\,{}^*g^{ij}\sigma\left(\frac{1}{\sigma}\right)_{;ij}, \end{cases} $$

where σ is defined by $e\sigma^2 = {}^*g^{mm} = 1/{}^*g_{mm}$, $e = \pm 1$; $\rho_m = \partial_m \log \rho$, $\sigma_i = \partial_i \log \sigma$; ${}'R_{ij}$ is the Ricci tensor of ${}^*g_{ij} = \rho^{-2} g_{ij}$; and the semi-colon followed by the indices i and j or m denotes covariant differentiation with respect to ${}^*g_{ij}$ or ${}^*g_{mm}$, respectively.[10] The condition for ${}^*g_{\alpha\beta}$ to represent an E_m is

$$ (2.3) \qquad {}^*R_{\alpha\beta} = -(m-1)c\,{}^*g_{\alpha\beta}, $$

where c is the constant scalar curvature of E_m.

We now suppose that $\partial_m \rho \neq 0$, i.e. that the totally umbilical hypersurfaces $x^m = \text{const.}$ are not totally geodesic in E_m. From (2.3) and $(2.2)_1$ we have

$$ \partial_i \rho_m + \rho_m \sigma_i = 0, $$

which gives on integration

$$ (2.4) \qquad \sigma \rho_m = z(x^m) \neq 0, $$

where $z(x^m)$ is some function of x^m alone. If we transvect $(2.2)_2$ and $(2.2)_3$ with ${}^*g^{ij}$ and ${}^*g^{mm}$ respectively and subtract the results, we find, on making use of (2.3) and (2.4),

$$ (2.5) \qquad -\frac{{}^*g^{ij}\,{}'R_{ij}}{n(n-1)} = c + e[z(x^m)]^2 \equiv a(x^m). $$

Since by definition the left-hand member is the scalar curvature of ${}^*g_{ij}$, this equation shows that *for each value of x^m the fundamental tensor ${}^*g_{ij}$ has constant scalar curvature*, which we denote by $a(x^m)$. We observe that since $z(x^m) \neq 0$, $a(x^m)$ cannot be identically equal to c.

[10] In this operation, the x^m in ${}^*g_{ij}$ and the x^k in ${}^*g_{mm}$, respectively, are considered as parameters. The covariant derivative with respect to ${}^*g_{mm}$ is formed with the Christoffel symbol of the second kind ${}'\Gamma^m_{mm}$ for ${}^*g_{mm}$.

Now equations (2.4) and (2.5) can be solved for $e\sigma^{-2}$, giving

$$e\sigma^{-2} = \frac{(\rho_m)^2}{a(x^m) - c} = \frac{(\partial_m \log \rho)^2}{a(x^m) - c}. \tag{2.6}$$

Therefore the first part of our theorem will be proved if we can show that in consequence of (2.4) and (2.5), equations (2.2) and (2.3) are equivalent to condition (3) of the theorem.

Because of (2.3), equation $(2.2)_2$ is of the form

$$\left(\frac{1}{\sigma}\right)_{;ij} + \frac{1}{\sigma}\,'R_{ij} = \tau\,{}^*g_{ij}\,, \tag{2.7}$$

where τ is a certain scalar. The integrability condition of this equation is

$$'R^l_{ijk}\left(\frac{1}{\sigma}\right)_{;l} + ('R_{ij;k} - 'R_{ik;j})\frac{1}{\sigma} + 'R_{ij}\left(\frac{1}{\sigma}\right)_{;k} - 'R_{ik}\left(\frac{1}{\sigma}\right)_{;j} = g_{ij}\tau_{;k} - g_{ik}\tau_{;j}\,,$$

where $'R^l_{ijk}$ is the Riemann tensor of ${}^*g_{ij}$. When we transvect this with ${}^*g^{ij}$ and then make use of the well-known identity (Eisenhart **5,** 82, (26.4))

$$'R^j_{k;j} = \tfrac{1}{2}\,'R_{;k}\,, \quad \text{where} \quad 'R = {}^*g^{ij}\,'R_{ij}\,,$$

we find

$$\frac{1}{2\sigma}\,'R_{;k} + 'R\left(\frac{1}{\sigma}\right)_{;k} = (n-1)\tau_{;k}\,.$$

But ${}^*g_{ij}$ is of constant scalar curvature $a(x^m)$ and $'R = -n(n-1)a(x^m)$. Therefore the above equation becomes $-na(x^m)\left(\frac{1}{\sigma}\right)_{;k} = \tau_{;k}$, which gives, on integration,

$$\tau = -\left[n\,\frac{a(x^m)}{\sigma} + w(x^m)\right],$$

where $w(x^m)$ is some function of x^m alone. With this value for τ, equation (2.7) can now be written more precisely as

$$\left(\frac{1}{\sigma}\right)_{;ij} + \frac{1}{\sigma}\,'R_{ij} = -\left[n\,\frac{a(x^m)}{\sigma} + w(x^m)\right]{}^*g_{ij}\,. \tag{2.8}$$

When this expression for $\left(\frac{1}{\sigma}\right)_{;ij}$ is substituted in $(2.2)_3$, the latter becomes, because of (2.3) and (2.5),

$$e\sigma^2\rho\left(\frac{1}{\rho}\right)_{;mm} + c - a(x^m) = \sigma w(x^m). \tag{2.9}$$

Now if ${}'\Gamma^m_{mm}$ is the Christoffel symbol of the second kind for ${}^*g_{mm}$, then

$$\left(\frac{1}{\rho}\right)_{;mm} = \partial_m\left(-\frac{\rho_m}{\rho}\right) - {}'\Gamma^m_{mm}\left(-\frac{\rho_m}{\rho}\right)$$

$$= \frac{\rho_m}{\rho}\left(-\partial_m \log\frac{\rho_m}{\rho} + \tfrac{1}{2}\partial_m \log {}^*g_{mm}\right)$$

$$= \frac{\rho_m}{\rho}\,\partial_m \log\frac{\rho}{\sigma\rho_m} = \frac{\rho_m}{\rho}\left[\rho_m - \frac{z'(x^m)}{z(x^m)}\right]$$

by (2.4).[11] Substituting this in (2.9) and using (2.4) and (2.5), we have

$$w(x^m) = -ez'(x^m) = -\frac{a'(x^m)}{2z(x^m)} = -\frac{a'(x^m)}{2\sigma\rho_m}. \tag{2.10}$$

Therefore (2.8) becomes

$$\left(\frac{1}{\sigma}\right)_{;ij} + \frac{1}{\sigma}\,{}'R_{ij} = -\frac{1}{\sigma}\left[na(x^m) - \frac{a'(x^m)}{2\rho_m}\right]{}^*g_{ij}. \tag{2.11}$$

It is easily seen from the way (2.5) was derived from $(2.2)_2$ and $(2.2)_3$, that on account of (2.4) and (2.5), equations (2.2) and (2.3) are equivalent to (2.11).

We now suppose, without loss of generality, that g_{ij} is a tensor of the family ${}^*g_{ij}$ so that[12]

$$\rho(x^k, x^m_o) = 1 \tag{2.12}$$

for some value x^m_o of x^m. We wish to write (2.11) in terms of ρ and g_{ij}. Connecting the covariant derivatives, Ricci tensors, and scalar curvatures related to the fundamental tensors ${}^*g_{ij} = \rho^{-2}g_{ij}$ and g_{ij} are the following formulas (Eisenhart **5**, 89–90):

$$\begin{cases} \left(\dfrac{1}{\sigma}\right)_{;ij} = \dfrac{1}{\rho}\left(\dfrac{\rho}{\sigma}\right)_{,ij} - \dfrac{\rho_{,ij}}{\sigma} - g^{hk}\left(\dfrac{1}{\sigma}\right)_{,h}\rho_{,k}g_{ij}, \\ {}'R_{ij} = R_{ij} - (n-2)\,\dfrac{\rho_{,ij}}{\rho} + g_{ij}g^{hk}\left[-\dfrac{\rho_{,hk}}{\rho} + (n-1)\,\dfrac{\rho_{,h}\rho_{,k}}{\rho^2}\right], \\ a(x^m) = a(x^m_o)\rho^2 + \dfrac{2}{n}\,\rho g^{hk}\rho_{,hk} - g^{hk}\rho_{,h}\rho_{,k}, \end{cases} \tag{2.13}$$

where as usual the comma denotes covariant derivative with respect to g_{ij}. We can now readily verify that on substitution from (2.13) and making use of (2.4), equation (2.11) becomes

$$\frac{1}{\rho}\,(\partial_m\rho)_{,ij} - (n-1)\,\frac{\partial_m\rho}{\rho^2}\,\rho_{,ij} + \rho_m R_{ij}$$

$$= \left[\frac{g^{hk}\partial_m(\rho_{,hk})}{n\rho} - \frac{(n-1)}{n}\,\frac{\partial_m\rho}{\rho^2}\,g^{hk}\rho_{,hk} - (n-1)a(x^m_o)\rho_m\right]g_{ij}. \tag{2.14}$$

[11] A prime (on the right) always means differentiation.

[12] See footnote 9.

Since $(\partial_m\rho)_{,ij} = \partial_m(\rho_{,ij})$, the above equation gives an identity when transvected by g^{ij} and therefore is equivalent to

$$\frac{1}{\rho}\,\partial_m(\rho_{,ij}) - (n-1)\frac{\partial_m\rho}{\rho^2}\,\rho_{,ij} + \rho_m R_{ij} \backsim g_{ij}\,,$$

i.e.

$$\partial_m\left(\frac{\rho_{,ij}}{\rho^{n-1}}\right) + \frac{\partial_m\rho}{\rho^{n-1}}\,R_{ij} \backsim g_{ij}\,, \tag{2.15}$$

Equation (2.15) proves condition (3) of our theorem.

Suppose $n > 2$. Integrating (2.15) with respect to x^m, we have

$$\frac{\rho_{,ij}}{\rho^{n-1}} - \frac{R_{ij}}{(n-2)\rho^{n-2}} + T_{ij} \backsim g_{ij}\,, \tag{2.16}$$

where T_{ij} is an integration tensor independent of x^m. Now in consequence of (2.12), we have

$$(\rho_{,ij})_{x^m=x_0^m} = 0,$$

as follows from the very definition of partial differentiation. Therefore if we put $x^m = x_o^m$ in (2.16), the result is

$$-\frac{R_{ij}}{n-2} + T_{ij} \backsim g_{ij}\,.$$

Using this expression for T_{ij} in (2.16), we have

$$\rho_{,ij} + \frac{\rho^{n-1} - \rho}{n-2}\,R_{ij} \backsim g_{ij}\,, \tag{2.17}$$

which is equivalent to (2.15). Equation (2.15) is also equivalent to

$$'R_{ij} - \rho^{n-2}R_{ij} \sim g_{ij}\,, \tag{2.18}$$

as follows by elimination of $\rho_{,ij}$ from $(2.13)_2$ and (2.17).

Finally, if $n = 2$, the equation $R_{ij} \sim g_{ij}$ is satisfied, and therefore (2.15) becomes

$$\partial_m\left(\frac{\rho_{,ij}}{\rho}\right) \backsim g_{ij}\,.$$

If we integrate this with respect to x^m and treat the result as we did (2.16), we can show that (2.15) is equivalent to

$$\rho_{,ij} \sim g_{ij}\,. \tag{2.19}$$

Equations (2.17), (2.18) and (2.19) complete the proof of our theorem.

We end this section by obtaining a direct consequence of Theorem 2.1. Let $n > 2$. If $R_{ij} = -(n-1)a(x_o^m)g_{ij}$, then (2.18) demands that $'R_{ij} = -(n-1)a(x^m)^*g_{ij}$. Also, if ρ is a function of x^m alone so that $'R_{ij} = R_{ij}$,

equation (2.18) becomes ${}'R_{ij} = -(n-1)a(x^m){}^*g_{ij}$. Hence if we remember that any tensor of the family ${}^*g_{ij}$ may be made to play the part of g_{ij}, we have

THEOREM 2.2. *Let an $E_{n+1}(n > 2)$ admit a one-parameter family of totally umbilical, but not totally geodesic, hypersurfaces V_n. Then the V_n are all E_n, if one of them is an E_n or if every V_n is trivially conformal*[13] *to another. Conversely, let c be a constant and $[\rho(x^k, x^m)]^{-2}g_{ij}(x^k)$ with x^m as parameter be the fundamental tensors of ∞^1 conformal $E_n(n > 2)$ whose scalar curvatures are not all equal to c. Then if $\partial_m\rho \neq 0$, there exists a unique E_{n+1} of scalar curvature c in which the given E_n can be imbedded as totally umbilical hypersurfaces.*

The latter part of this theorem has been mentioned at the beginning of this paper.

3.1 An imbedding problem

Suppose that for a V_n the equation

$$\begin{cases} \rho_{,ij} + \dfrac{\rho^{n-1} - \rho}{n-2} R_{ij} \backsim g_{ij} & \text{for } n > 2, \\ \rho_{,ij} \backsim g_{ij} & \text{for } n = 2 \end{cases} \tag{3.1}$$

admits a solution for ρ. We now prove that if the scalar curvature of V_n is constant, then that of $\rho^{-2}g_{ij}$ is also constant. This result combined with Theorem 2.1 gives the following

THEOREM 3.1. *A necessary and sufficient condition for a V_n with fundamental tensor g_{ij} to be imbedable in an E_{n+1} as a member of ∞^1 totally umbilical, but not totally geodesic, hypersurfaces is that V_n has constant scalar curvature and equation* (3.1) *admits a one-parameter family of solutions $\rho(x^k, x^m)$ which is such that $\rho(x^k, x_o^m) = 1$ for some value x_o^m of the parameter x^m. If a V_n satisfies this condition and ρ is any one-parameter family of solutions of* (3.1) *of the above-mentioned nature, then $\rho^{-2}g_{ij}$ are fundamental tensors of a family of V_n which can be imbedded in an E_{n+1} as totally umbilical hypersurfaces.*

PROOF. Suppose that the scalar curvature $a = -(R/n(n-1))$ of V_n is constant. Consider first the case $n > 2$. We write $(3.1)_1$ as

$$\rho_{,ij} + \phi R_{ij} = \psi g_{ij}, \tag{3.1'}_1$$

where ψ is an unknown scalar and

$$\phi = \frac{\rho^{n-1} - \rho}{n-2}. \tag{3.2}$$

The integrability condition of $(3.1')_1$ is

$$R^l_{ijk}\rho_{,l} + \phi' R_{i[j}\rho_{,k]} + \phi R_{i[j,k]} = g_{i[j}\psi_{,k]}.$$

[13] Two V_n are said to be *trivially conformal* to each other, if their fundamental tensors (in the same coordinates) differ by a constant scalar factor.

Transvection of this with g^{ij} gives, because $R_{i[j,k]}g^{ij} = \frac{1}{2}R_{,k} = 0$ (Eisenhart **5**, 82, (26.4)),

$$R^l_k\rho_{,l} = \frac{n-1}{1-\phi'}(\psi_{,k} + na\phi'\rho_{,k}). \tag{3.3}$$

Using $(3.1')_1$, (3.2) and (3.3), we have

$$\begin{aligned}(g^{hk}\rho_{,h}\rho_{,k})_{,i} &= 2g^{hk}\rho_{,h}\rho_{,ki}\\ &= 2[(\rho\psi)_{,i} + na\phi'\rho\rho_{,i}].\end{aligned}$$

Let us then substitute for ϕ' from (3.2) and integrate; this gives

$$g^{hk}\rho_{,h}\rho_{,k} = 2\rho\psi + \frac{2(n-1)a\rho^n - na\rho^2}{n-2} - A, \qquad A = \text{const.} \tag{3.4}$$

Now the scalar curvature $'a$ of $\rho^{-2}g_{ij}$ is (cf. (2.13))

$$'a = a\rho^2 + \tfrac{1}{2}n\rho g^{hk}\rho_{,hk} - g^{hk}\rho_{,h}\rho_{,k},$$

and it is readily shown that on account of $(3.1')_1$, (3.2) and (3.4), the right-hand member of the above equation is equal to A. Hence the scalar curvature of $\rho^{-2}g_{ij}$ is constant, and our theorem for the case $n > 2$ is proved.

Equation $(3.1)_2$ for the case $n = 2$ can be treated in like manner. Indeed, we have in this case

$$\begin{aligned}\rho_{,ij} &= \psi g_{ij}, \qquad & R^l_k\rho_{,l} &= \psi_{,k},\\ R_{hk} &= -ag_{hk}, \qquad & \psi &= -(a\rho + B),\end{aligned}$$

where B is a constant. The first two of these equations correspond to $(3.1')_1$ and (3.3) respectively; the third holds because $n = 2$; and the last follows at once from the second and third equations. We can then easily verify that

$$g^{hk}\rho_{,h}\rho_{,k} = -(a\rho^2 + 2B\rho + A), \qquad A = \text{const.},$$

and hence that A is the scalar curvature of $\rho^{-2}g_{ij}$. Our theorem is thus completely proved.

3.2. Continuation

The question naturally arises: Given a V_n with constant scalar curvature, how can we know whether or not it is imbedable in an E_{n+1} as a member of ∞^1 totally umbilical hypersurfaces? The answer given in Theorem 3.1 is not explicit enough, but may be so reduced that it depends only on the compatibility of a system of polynomial equations.

Consider first the general case $n > 2$. It is readily seen from $(3.1')_1$, (3.2) and (3.3) that equation $(3.1)_1$ is equivalent to the following system of linear differential equations in the $n + 2$ unknown functions ρ, ψ, θ_i:

$$\begin{cases}\theta_{i,j} + \phi(\rho)R_{ij} = \psi g_{ij},\\ \psi_{,i} = \dfrac{1-\phi'}{n-1}R^l_i\theta_l - na\phi'\theta_i,\\ \rho_{,i} = \theta_i.\end{cases} \tag{3.5}$$

We now apply a classical theorem (Eisenhart **6,** 1–4) on a system of linear differential equations to (3.5). The integrability conditions of $(3.5)_1$, $(3.5)_2$, $(3.5)_3$ are respectively

$$(3.6)\quad \begin{cases} R^l_{ijk}\theta_l + \phi' R_{i[j}\rho_{,k]} + \phi R_{i[j,k]} = g_{i[j}\psi_{,k]}, \\ (1-\phi')\theta_l R^l_{[i,j]} = (1-\phi')\theta_l R^l_{[i,j]} + (1-\phi')\theta_{l,[j}R^l_{i]} - \phi''\theta_l R^l_{[i}\rho_{,j]} = 0, \\ \theta_{[i,j]} = 0. \end{cases}$$

Equation $(3.6)_3$ is satisfied on account of $(3.5)_1$. On substitution from (3.5), equations $(3.6)_1$ and $(3.6)_2$ become

$$(3.7)\quad \begin{cases} \theta_l\left[R^l_{ijk} - \frac{1}{n-1}g_{i[j}R^i_{k]} + \phi'\left(R_{i[j}\delta^l_{k]} + \frac{1}{n-1}g_{i[j}R^l_{k]} + na\,g_{i[j}\delta^l_{k]}\right)\right] \\ \qquad\qquad + \phi R_{i[j,k]} = 0, \\ (1-\phi')\theta_l R^l_{[i,j]} - \phi''\theta_l\theta_h R^l_{[i}\delta^h_{j]} = 0. \end{cases}$$

It is easily verified that $(3.7)_2$ is identical with $(3.7)_1$ transvected by θ^i. This was also to be expected, because (3.3) was obtained from the integrability condition of $(3.1')_1$. Hence the integrability condition of (3.5) reduces to $(3.7)_1$, which we write as

$$(F)_1 \qquad (A^l_{ijk} + B^l_{ijk}\phi')\theta_l + C_{ijk}\phi = 0,$$

where

$$(3.8)\quad \begin{cases} A^l_{ijk} = R^l_{ijk} - \frac{1}{n-1}g_{i[j}R^l_{k]}, \\ B^l_{ijk} = R_{i[j}\delta^l_{k]} + \frac{1}{n-1}g_{i[j}R^l_{k]} + nag_{i[j}\delta^l_{k]}, \\ C_{ijk} = R_{i[j,k]} \end{cases}$$

are tensors constructed from g_{ij}.

It is not difficult to show that $(F)_1$ is identically satisfied if and only if V_n is an S_n. If $(F)_1$ is not identically satisfied, we differentiate it covariantly with respect to g_{ij} and substitute for $\theta_{i,j}$, $\psi_{,i}$ and $\rho_{,i}$ from (3.5). The result is

$$(F)_2 \qquad \begin{aligned} &\phi'' B^l_{ijk}\theta_l\theta_h + [A^l_{ijk,h} + (B^l_{ijk,h} + C_{ijk}\delta^l_h)\phi']\theta_l \\ &\qquad + (A^l_{ijk} + B^l_{ijk}\phi')(-\phi R_{lh} + \psi g_{lh}) + C_{ijk,l}\phi = 0. \end{aligned}$$

Because of $(F)_1$, this cannot be identically satisfied without $(F)_1$ first being identically satisfied. Differentiating $(F)_2$ covariantly and then substituting for $\theta_{i,j}$, $\psi_{,i}$ and $\rho_{,i}$ from (3.5), we get a set of equations $(F)_3$. Proceeding in this way we obtain a sequence of sets $(F)_1$, $(F)_2$, etc. of equations, to which a classical theorem on mixed systems of total differential equations (Eisenhart **6,** 3) can be directly applied. We observe that the equations $(F)_1$, $(F)_2$, etc. in our

case are all polynomial equations and that they can be constructed as soon as the fundamental tensor g_{ij} of V_n is given.

Hence combining the preceding results with Theorem 3.1 we have the necessity of the conditions in the following

THEOREM 3.2. *In order that a $V_n (n > 2)$ may be imbedded in an E_{n+1} as a member of ∞^1 totally umbilical, but not totally geodesic, hypersurfaces, it is necessary and sufficient that the following conditions are satisfied:*

(1) *A positive integer N exists such that the equations of the sets $(F)_1$, $(F)_2$, $\cdots$, $(F)_N$ admit a one-parameter family of solutions $\rho(x^k, x^m)$, $\psi(x^k, x^m)$, $\theta_i(x^k, x^m)$ of the nature that $\rho(x^k, x_o^m) = 1$ for some value x_o^m of the parameter x^m;*

(2) *The equations of the set $(F)_{N+1}$ are satisfied because of the equations of the sets $(F)_1$, $(F)_2$, $\cdots$, $(F)_N$.*

To prove the sufficiency of the conditions we need only show that a V_n satisfying them must be of constant scalar curvature. But the latter fact is easily proved; for, if we transvect $(F)_1$ with g^{ij}, the result is $\frac{1}{2}\phi R_{,k} = 0$, showing that R is constant.

A V_2 with constant scalar curvature is an S_2. For an S_n equation (3.1) becomes

$$\rho_{,ij} \backsim g_{ij}. \tag{3.9}$$

If we use the Riemann form of the fundamental form of S_n (Eisenhart **5,** 85)

$$g_{ij}\,dx^i\,dx^j = \frac{e_1\,(dx^1)^2 + \cdots + e_n\,(dx^n)^2}{\left\{1 + \frac{a}{4}\,[e_1(x^1)^2 + \cdots + e_n(x^n)^2]\right\}^2} \qquad (e_i = \pm 1), \tag{3.10}$$

where a is the (constant) scalar curvature of S_n, then the general solution of (3.9) is readily seen to be

$$\rho = \frac{A_o[e_1(x^1)^2 + \cdots + e_n(x^n)^2] + A_i x^i + A_{n+1}}{1 + \frac{a}{4}\,[e_1(x^1)^2 + \cdots + e_n(x^n)^2]}, \tag{3.11}$$

where the A's are arbitrary constants Evidently, by suitable choice of the A's as functions of a parameter x^m, a one-parameter family of solutions $\rho(x^k, x^m)$ can be found (and in infinitely many ways) such that $\rho(x^k, x_o^m) = 1$ for some value x_o^m of x^m. Hence the condition of Theorem 3.1 is satisfied. Now it is seen from (3.10) and (3.11) that in this case every tensor of the family $\rho^{-2}g_{ij}$ is the fundamental tensor of an S_n. And since it is known (Wong **14,** Th. 9.3) that an E_{n+1} which admits $\infty^1 S_n$ as totally umbilical hypersurfaces is an S_{n+1}, we have

THEOREM 3.3. *An S_n can always be imbedded (and in infinitely many ways) in an S_{n+1} as a member of ∞^1 totally umbilical hypersurfaces.*

4. Totally geodesic hypersurfaces in an E_{n+1}

We now consider the case excluded from the preceding discussions. As follows from what we said at the beginning of §2, the condition for a V_n with

fundamental tensor g_{ij} to be imbedable in an E_m ($m = n + 1$) as a member of ∞^1 totally geodesic hypersurfaces is equivalent to the condition for the existance of a non-zero scalar $\rho(x^\alpha)$ such that the fundamental tensor ${}^*g_{\alpha\beta}$ with components

$$(4.1) \qquad {}^*g_{im} = 0, \qquad {}^*g_{ij} = g_{ij}(x^k), \qquad {}^*g_{mm} = e\rho^2, \qquad (e = \pm 1),$$

represents an E_m. By definition, ${}^*g_{\alpha\beta}$ represents an E_m if and only if

$$(4.2) \qquad {}^*R_{\alpha\beta} = -nc\,{}^*g_{\alpha\beta}, \qquad c = \text{const.}$$

Now the components of the Ricci tensor of ${}^*g_{\alpha\beta}$ are

$$(4.3) \qquad {}^*R_{im} = 0, \qquad {}^*R_{ij} = R_{ij} + \frac{\rho_{,ij}}{\rho}, \qquad {}^*R_{mm} = {}^*g_{mm}g^{ij}\frac{\rho_{,ij}}{\rho}.$$

Therefore (4.2) becomes, on account of (4.1),

$$(4.4) \qquad \rho_{,ij} + \rho R_{ij} = -nc\rho g_{ij}, \qquad g^{ij}\rho_{,ij} = -nc\rho.$$

Transvecting $(4.4)_1$ with g^{ij} and comparing the result with $(4.4)_2$, we find

$$(4.5) \qquad R = g^{ij}R_{ij} = -n(n-1)c.$$

This shows that the scalar curvature a of V_n is equal to the (constant) scalar curvature c of E_m. Hence from $(4.4)_1$ we have the following theorem complementary to Theorem 3.1 of the preceding section:

THEOREM 4.1. *A necessary and sufficient condition for a V_n with fundamental tensor g_{ij} to be imbedable in an E_{n+1} as a member of ∞^1 totally geodesic hypersurfaces is that the scalar curvature a of V_n is constant and the equation*

$$(4.6) \qquad \rho_{,ij} + \rho R_{ij} = -na\rho g_{ij}$$

admits a non-zero solution for ρ. If a V_n satisfies this condition, then corresponding to each non-zero solution ρ of (4.6) *there are two and only two such imbedding E_{n+1}, and the components of their fundamental tensors ${}^*g_{\alpha\beta}$ are given by* (4.1).

Equation (4.6) is equivalent to the following system of linear differential equations in the $n + 1$ unknown functions ρ and θ_i:

$$(4.7) \qquad \theta_{i,j} + \rho R_{ij} = -na\rho g_{ij}, \qquad \rho_{,i} = \theta_i,$$

for which the sets $(G)_1$, $(G)_2$, etc. of equations corresponding to $(F)_1$, $(F)_2$, etc. of §3.2 are

$$(G_1) \qquad (A^l_{ijk} + B^l_{ijk})\theta_l + C_{ijk}\rho = 0,$$

$$(G_2) \qquad (A^l_{ijk,h} + B^l_{ijk,h} + C_{ijk}\delta^l_h)\theta_l - [(A^l_{ijk} + B^l_{ijk})(R_{lh} + nag_{lh}) - C_{ijk,h}]\rho = 0,$$

etc.,

where the tensors A, B and C are as defined in (3.8). Since the equations in $(G)_1$, $(G)_2$, etc. are all linear homogeneous in ρ and θ_i, we have from Theorem 4.1 the following theorem corresponding to Theorem 3.2:

THEOREM 4.2. *In order that a V_n may be imbedded in an E_{n+1} as a member of ∞^1 totally geodesic hypersurfaces, it is necessary and sufficient that the following conditions are satisfied*:

(1) *A positive integer N exists such that the sets of equations $(G)_1, (G)_2, \cdots, (G)_N$ are compatible and the equations of the sets $(G)_{N+1}$ are satisfied because of the former sets*;

(2) *The sets $(G)_1, (G)_2, \cdots, (G)_N$ have at most n independent equations.*

For an S_n, equations (4.7) become

$$\theta_{i,j} = -a\rho g_{ij}, \qquad \rho_{,i} = \theta_i, \tag{4.8}$$

which are completely integrable. Hence by the remark just above Theorem 3.3, we have

THEOREM 4.3. *An S_n can always be imbedded (and in infinitely many ways) in an S_{n+1} as a member of ∞^1 totally geodesic hypersurfaces.*

The case of an S_n is closed by Theorems 3.3 and 4.3, and will therefore be excluded from our future discussions.

5. Totally umbilical hypersurfaces in an E_{n+1} which are C_n

Equations (3.1) and (4.6) which appear in Theorems 3.1 and 4.1 are all particular cases of equation (6.1) of the next section, where the latter equation will be studied in detail for the case of a C_n, i.e. a conformal-Euclidean V_n. In light of what we shall obtain, the following theorems are immediate consequences of Theorems 3.1 and 4.1:

THEOREM 5.1. *If an $E_{n+1}(n > 2)$ admits a family of totally umbilical hypersurfaces which are C_n but not S_n, then the tensor L_{ij} of each of these hypersurfaces is of the form* (6.3) *or* (6.4). *Conversely, a C_n of constant scalar curvature with tensor L_{ij} of the form* (6.3) *or* (6.4) *can be imbedded in infinitely many E_{n+1} as a member of ∞^1 totally umbilical hypersurfaces or totally geodesic hypersurfaces. The determination of all these E_{n+1} is based either on a simple integration or on the solution of a linear differential equation of the second order.* (Cf. Ths. 6.1, 7.1–7.3)

THEOREM 5.2. *A C_n can be imbedded in an E_{n+1} as a member of ∞^1 totally umbilical hypersurfaces or totally geodesic hypersurfaces, if and only if it is an S_n or if $n > 2$ and its fundamental form can be reduced to the form* (8.8′), (8.9′) *or* (8.27). (Cf. Ths. 8.1, 8.2.)

II. SOME CONFORMAL-EUCLIDEAN SPACES C_n

6.1. A differential equation

For the rest of this paper we confine ourselves to the case of a $C_n(n > 2)$, and study in detail the equation

$$\rho_{,ij} + \phi L_{ij} = \omega g_{ij}, \tag{6.1}$$

where L_{ij} is the tensor (1.2), ω is an unspecified scalar, and $\phi = \phi(\rho)$ is a given function of ρ such that $\phi' \neq -1/(n-2)$.

The supposition $\phi' \neq -1/(n-2)$ is significant, because for any C_n the equation

$$\rho_{,ij} - \frac{\rho + A}{n-2} L_{ij} = \omega g_{ij}, \qquad A = \text{const.} \tag{6.2}$$

always has a solution. Indeed, if g_{ij} represents a C_n, a scalar σ exists such that the tensor $\sigma^{-2}g_{ij}$ represents a Euclidean space. Consequently σ satisfies the equation

$$R_{ij} - (n-2)\frac{\sigma_{,ij}}{\sigma} + g_{ij}\left[-\frac{g^{hk}\sigma_{,hk}}{\sigma} + (n-1)\frac{g^{hk}\sigma_{,h}\sigma_{,k}}{\sigma^2}\right] = 0,$$

the left-hand member being the Ricci tensor of $\sigma^{-2}g_{ij}$ written in terms of g_{ij} (cf. $(2.13)_2$). From this it follows that $\rho = \sigma - A$ is a solution of (6.2).

We now proceed to prove the following

THEOREM 6.1. *For a C_n, which is not an S_n, equation* (6.1) *has a solution for ρ, if and only if the tensor L_{ij} of C_n is of the form*[14]

$$L_{ij} = \bar{\eta}(\bar{\rho})\bar{\rho}_{,i}\bar{\rho}_{,j} + \bar{\xi}(\bar{\rho})g_{ij}, \qquad g_{ij}\bar{\rho}_{,i}\bar{\rho}_{,j} \neq 0, \tag{6.3}$$

or

$$L_{ij} = \bar{\eta}(\bar{\rho})\bar{\rho}_{,i}\bar{\rho}_{,j}, \qquad g^{ij}\bar{\rho}_{,i}\bar{\rho}_{,j} = 0, \tag{6.4}$$

where $\bar{\eta}(\bar{\rho}) \neq 0$, $\bar{\xi}(\bar{\rho}) \neq 0$; in the latter case $L = g^{ij}L_{ij} = 0$, $\omega = 0$. Let a C_n satisfy either of these conditions. Then a function $\bar{\lambda} = \bar{\lambda}(\bar{\rho})$ exists such that

$$\bar{\rho}_{,ij} = \bar{\lambda}\bar{\rho}_{,i}\bar{\rho}_{,j} + \frac{\bar{\xi}'}{\bar{\eta}}g_{ij}, \quad \text{or} \quad \bar{\rho}_{,ij} = \bar{\lambda}\bar{\rho}_{,i}\bar{\rho}_{,j}, \tag{6.5}$$

according as we have (6.3) *or* (6.4); *and the solution of* (6.1) *is identical with the solution $\rho = \rho(\bar{\rho})$ of the linear differential equation*

$$\rho'' + \bar{\lambda}\rho' + \phi\bar{\eta} = 0. \tag{6.6}$$

PROOF. For a C_n, equations (1.4) hold. The integrability condition of (6.1) is

$$R^l_{ijk}\rho_{,l} + \phi' L_{i[j}\rho_{,k]} + \phi L_{i[j,k]} = g_{i[j}\omega_{,k]},$$

which, on account of (1.4), becomes

$$\frac{1}{n-2}(L_{i[j}\rho_{,k]} + g_{i[j}L^l_{k]}\rho_{,l}) + \phi' L_{i[j}\rho_{,k]} = g_{i[j}\omega_{,k]},$$

i.e.

$$\left(\phi' + \frac{1}{n-2}\right)L_{i[j}\rho_{,k]} + g_{i[j}\left(\frac{1}{n-2}L^l_{k]}\rho_{,l} - \omega_{,k]}\right) = 0. \tag{6.7}$$

[14] That V_n be a C_n satisfying (6.3) is, incidentally, the condition for a V_n ($n > 2$), not an S_n, to be a *subprojective space* in the sense of Kagan. Cf. Schouten-Struik **11**, 215–225, and also the literatures quoted there.

Transvection of this with g^{ij} gives

$$(1-\phi')L^l_k\rho_{,l} - (n-1)\omega_{,k} = -\left(\phi' + \frac{1}{n-2}\right)L\rho_{,k}\,. \tag{6.8}$$

Transvecting (6.7) again with $g^{hi}\rho_{,h}$, we find

$$\phi'\rho_{,l}L^l_{[j}\rho_{,k]} + \omega_{,[j}\rho_{,k]} = 0,$$

which is equivalent to

$$\phi' L^l_k\rho_{,l} + \omega_{,k} = \alpha\rho_{,k}\,, \qquad \alpha = \text{scalar}. \tag{6.9}$$

Since $\phi' \neq -1/(n-2)$, equations (6.8) and (6.9) can be solved for $L^l_k\rho_{,l}$ and $\omega_{,k}$, giving

$$L^l_k\rho_{,l} = \beta\rho_{,k}\,, \qquad \omega_{,k} = \gamma\rho_{,k}\,, \tag{6.10}$$

where β and γ are some scalars. Equation $(6.10)_2$ shows that $\omega = \omega(\rho)$ and $\gamma = \omega'$. Consequently, as a result of (6.10), the integrability condition (6.7) becomes

$$\left(\phi' + \frac{1}{n-2}\right)L_{i[j}\rho_{,k]} + \left(\frac{1}{n-2}\beta - \omega'\right)g_{i[j}\rho_{,k]} = 0, \tag{6.11}$$

which can be written in the form

$$L_{ij} = \eta\rho_{,i}\rho_{,j} + \zeta g_{ij}\,. \tag{6.12}$$

We shall now proceed to prove that *η and ζ are both functions of ρ alone.*

On account of (6.12) we have

$$L = \xi\eta + n\zeta, \tag{6.13}$$

$$[2 + (n-2)\phi']\zeta = (n-2)\omega' - \xi\eta, \tag{6.14}$$

where

$$\xi \equiv g^{ij}\rho_{,i}\rho_{,j}\,. \tag{6.15}$$

Equation (6.13) is the result of transvecting (6.12) with g^{ij}. To prove (6.14), we use (6.12) in (6.8) and obtain

$$(1-\phi')(\xi\eta + \zeta) - (n-1)\omega' = -\left(\phi' + \frac{1}{n-2}\right)L,$$

which gives (6.14) on substitution of L from (6.13).

Because of (6.12), equation (6.1) becomes

$$\rho_{,ij} = -\phi\eta\rho_{,i}\rho_{,j} + (\omega - \phi\zeta)g_{ij}\,. \tag{6.16}$$

Thus, by differentiating (6.15) we have

$$\begin{aligned}\xi_{,k} &= 2g^{ij}\rho_{,i}\rho_{,jk}\\ &= -2(\phi\xi\eta + \phi\zeta - \omega)\rho_{,k}\,.\end{aligned}$$

From this it follows that $\xi = \xi(\rho)$ and

$$(\xi\eta + \zeta)\phi - \omega = -\tfrac{1}{2}\xi'. \tag{6.17}$$

Using (6.12) in $(1.4)_2$, we have

$$g_{i[j}\zeta_{,k]} + \eta\rho_{,i[k}\rho_{,j]} + \rho_{,i}\rho_{,[j}\eta_{,k]} = 0.$$

On substitution from (6.16), this becomes

$$g_{i[j}\zeta_{,k]} - \eta(\omega - \phi\zeta)g_{i[j}\rho_{,k]} + \rho_{,i}\rho_{,[j}\eta_{,k]} = 0. \tag{6.18}$$

At this stage, we have to consider separately the two cases $\xi \neq 0$ and $\xi = 0$.

6.2. The case $\xi \neq 0$

Since ϕ and ω are functions of ρ, it follows either from (6.14) or from (6.17) that ζ is of the form

$$\zeta = \mu(\rho) + \nu(\rho)\eta, \tag{6.19}$$

where μ and ν are some functions of ρ, and $\nu \neq 0$ because $\xi \neq 0$. Using this value of ζ in (6.18), the latter becomes

$$(\nu g_{i[j} + \rho_{,i}\rho_{,[j})\eta_{,k]} + [\mu' + \nu'\eta - \eta(\omega - \phi\zeta)]g_{i[j}\rho_{,k]} = 0.$$

Since $\nu \neq 0$, this equation can be written

$$(\nu g_{i[j} + \rho_{,i}\rho_{,[j})(\eta_{,k]} - \tau\rho_{,k]}) = 0, \tag{6.20}$$

where

$$-\nu\tau = \mu' + \nu'\eta - \eta(\omega - \phi\zeta). \tag{6.21}$$

From (6.20) it follows that either

$$\eta_{,k} - \tau\rho_{,k} = 0, \tag{6.22}$$

or

$$\nu g_{ij} + \rho_{,i}\rho_{,j} = \epsilon(\eta_{,i} - \tau\rho_{,i})(\eta_{,j} - \tau\rho_{,j}), \qquad \epsilon = \text{a scalar.}$$

But the latter case cannot happen; otherwise g_{ij} would be of rank $\leqq 2$. Therefore (6.22) is true, and η is a function of ρ alone. Then it follows from (6.13) and (6.14) that ζ and L are also functions of ρ alone, as was to be proved.

Since τ is now equal to η', we have from (6.21) and (6.19) that

$$\zeta' = \eta(\omega - \phi\zeta). \tag{6.23}$$

If $\eta = 0$, it follows from (6.12) and (1.4) that C_n is an S_n. And from (6.14), (6.17), (6.23) and $\xi \neq 0$, we can easily show that if $\zeta = 0$, then $\eta = 0$ and consequently C_n is an S_n.

We have now proved the necessity of condition (6.3) for the case $\xi \neq 0$. That this same condition is also sufficient can be proved as follows.

Let a C_n be given satisfying (6.3). Then C_n cannot be an S_n. Using (6.3) in $(1.4)_2$, which is true for any C_n, we find

$$\bar{\eta}\bar{\rho}_{,i[k}\bar{\rho}_{,j]} + \bar{\zeta}'g_{i[j}\bar{\rho}_{,k]} = 0.$$

Because $\bar{\eta} \neq 0$, this can be written

$$\bar{\rho}_{,ij} = \bar{\lambda}\bar{\rho}_{,i}\bar{\rho}_{,j} + \frac{\bar{\xi}'}{\bar{\eta}} g_{ij}, \qquad \bar{\lambda} = \text{a scalar.} \tag{6.24}$$

Differentiating (6.15) and making use of (6.24), we have

$$\bar{\xi}_{,k} = 2g^{ij}\bar{\rho}_{,i}\bar{\rho}_{,jk} = 2g^{ij}\bar{\rho}_{,i}\left(\bar{\lambda}\bar{\rho}_{,j}\bar{\rho}_{,k} + \frac{\bar{\xi}'}{\bar{\eta}} g_{jk}\right) = 2\left(\bar{\lambda}\bar{\xi} + \frac{\bar{\xi}'}{\bar{\eta}}\right)\bar{\rho}_{,k}.$$

Since $\bar{\xi} \neq 0$, it follows from this that $\bar{\xi}$ and $\bar{\lambda}$ are functions of $\bar{\rho}$ alone and

$$\bar{\xi} = 2\left(\bar{\lambda}\bar{\xi} + \frac{\bar{\xi}'}{\bar{\eta}}\right). \tag{6.25}$$

If C_n, which is not an S_n, permits a solution of (6.1) such that $\xi = g^{ij}\rho_{,i}\rho_{,j} \neq 0$, equations (6.12) and (6.16) with $\eta \neq 0$, $\zeta \neq 0$ must hold. Comparison of $(6.3)_1$ with (6.12) tells us that $\rho_{,i}$ and $\bar{\rho}_{,i}$ have the same direction, which is a principal direction of L_{ij}. Hence $\bar{\rho} = \bar{\rho}(\rho)$, the exact form of which will presently be determined.

When $\bar{\rho} = \bar{\rho}(\rho)$ is used in $(6.3)_1$ and (6.24), the result is

$$L_{ij} = \eta\bar{\rho}'^2\rho_{,i}\rho_{,j} + \xi g_{ij}, \tag{6.26}$$

$$\bar{\rho}'\rho_{,ij} = (\bar{\lambda}\bar{\rho}'^2 - \bar{\rho}'')\rho_{,i}\rho_{,j} + \frac{\bar{\xi}'}{\bar{\eta}} g_{ij}. \tag{6.27}$$

Comparing these with (6.12) and (6.16) respectively, we find

$$\eta = \bar{\eta}\bar{\rho}'^2, \qquad \zeta = \bar{\xi}, \qquad \bar{\rho}' = \frac{\bar{\lambda}\bar{\rho}'^2 - \bar{\rho}''}{-\phi\eta} = \frac{\bar{\xi}'/\bar{\eta}}{\omega - \phi\zeta}.$$

From these it follows, by elimination of η and ζ, that

$$\phi = \frac{\bar{\rho}'' - \bar{\lambda}\bar{\rho}'^2}{\bar{\eta}\bar{\rho}'^3}, \tag{6.28}$$

$$\omega = \bar{\xi}\phi + \frac{\bar{\xi}'}{\bar{\eta}\bar{\rho}'} = \frac{(\bar{\rho}'' - \bar{\lambda}\bar{\rho}'^2)\bar{\xi} + \bar{\xi}'\bar{\rho}'^2}{\bar{\eta}\bar{\rho}'^3}. \tag{6.29}$$

Equation (6.28) is a relation between $\bar{\rho}$ and ρ, while (6.29) gives the value of the unspecified scalar ω. These equations must be satisfied if C_n permits a solution ρ of (6.1) such that $\xi \neq 0$. Now because of (6.28) and (6.29), equation (6.1) can be written

$$\bar{\eta}\bar{\rho}'^3\rho_{,ij} + (\bar{\rho}'' - \bar{\lambda}\bar{\rho}'^2)L_{ij} = [(\bar{\rho}'' - \bar{\lambda}\bar{\rho}'^2)\bar{\xi} + \bar{\xi}'\bar{\rho}'^2]g_{ij}.$$

But it will be noticed that this equation is identical with the result of eliminating $\rho_{,i}\rho_{,j}$ from (6.26) and (6.27) Hence, equation (6.1) is a consequence of (6.28), and we may conclude that for a C_n satisfying (6.3), the solution of (6.1) is identical with the inverted solution $\rho = \rho(\bar{\rho})$ of (6.28). But it is readily verified that written with ρ as dependent variable, (6.28) becomes (6.6). Hence our theorem for the case $\xi \neq 0$ is proved.

6.3. The case $\xi = 0$

In the preceding discussion the supposition $\xi \neq 0$ was first introduced after equation (6.18). Therefore this equation and the results preceding it also hold for the present case.

From (6.13), (6.14) and (6.17) we have, by putting $\xi = 0$,

$$L = n\zeta, \qquad [2 + (n-2)\phi']\zeta = (n-2)\omega' \qquad \omega = \phi\zeta. \tag{6.30}$$

Since ϕ and ω are known to be functions of ρ alone, it follows from the above equations that ζ *and* L *are also functions of* ρ *alone*. As a consequence of this and (6.30), equation (6.18) now becomes

$$(\zeta' g_{i[j} - \rho_{,i}\eta_{,[j})\rho_{,k]} = 0, \tag{6.31}$$

i.e.

$$\zeta' g_{ij} = \rho_{,i}\eta_{,j} + \alpha_i\rho_{,j} = \rho_{,j}\eta_{,i} + \alpha_j\rho_{,i}\,, \tag{6.31'}$$

where α_i is a certain vector. The last equation shows that

$$\alpha_i = \eta_{,i} + \beta\rho_{,i}\,, \qquad \beta = \text{a scalar},$$

and therefore (6.31′) becomes

$$\zeta' g_{ij} = \rho_{,i}\eta_{,j} + \rho_{,j}\eta_{,i} + \beta\rho_{,i}\rho_{,j}\,.$$

From this it follows that

$$\zeta' = 0; \tag{6.32}$$

otherwise g_{ij} would be of rank $\leqq 2$. Consequently, (6.31) reduces to $\eta_{,[j}\rho_{,k]} = 0$, which shows that η *is a function of* ρ *alone*.

Moreover, on account of (6.32), we have from (6.30) that

$$L = \zeta = \omega = 0. \tag{6.33}$$

Therefore, equations (6.12) and (6.1) become

$$L_{ij} = \eta(\rho)\rho_{,i}\rho_{,j}\,, \tag{6.34}$$

$$\rho_{,ij} = -\phi\eta\rho_{,i}\rho_{,j}\,, \tag{6.35}$$

where $\eta \neq 0$; otherwise C_n would be an S_n.

Equation (6.34) proves the necessity of condition (6.4) for the case $\xi = 0$. That this condition is also sufficient will now be proved by an argument similar to that for the case $\xi \neq 0$. Let there be given a C_n whose tensor L_{ij} is of the form (6.4). Using (6.4) in $(1.4)_2$, we deduce, as in (6.24),

$$\bar\rho_{,ij} = \bar\lambda\bar\rho_{,i}\bar\rho_{,j}\,, \qquad \bar\lambda = \text{a scalar}. \tag{6.36}$$

The integrability condition of this is

$$R^l_{i\,jk}\bar\rho_{,l} = \bar\lambda_{,[k}\bar\rho_{,j]}\bar\rho_{,i} + \bar\lambda\bar\rho_{,i[k}\bar\rho_{,j]}\,. \tag{6.37}$$

Now in consequence of $\xi = 0$, (6.4) and $(1.4)_1$, the left-hand member vanishes, while at the same time, the last term is zero because of (6.36). Hence (6.37) becomes $\bar{\lambda}_{,[k}\bar{\rho}_{,j]} = 0$, showing that $\bar{\lambda}$ *is a function of* $\bar{\rho}$ *alone*

Now by comparison of (6.34) and (6.4), it follows that $\bar{\rho} = \bar{\rho}(\rho)$. When this is substituted in (6.4) and (6.36), the result is

$$L_{ij} = \bar{\eta}\bar{\rho}'^2\rho_{,i}\rho_{,j}, \qquad \bar{\rho}'\rho_{,ij} = (\bar{\lambda}\bar{\rho}'^2 - \bar{\rho}'')\rho_{,i}\rho_{,j}. \tag{6.38}$$

Comparing these with (6.34) and (6.35), we have

$$\eta = \bar{\eta}\bar{\rho}'^2, \qquad \bar{\rho}' = \frac{\bar{\lambda}\bar{\rho}'^2 - \bar{\rho}''}{-\phi\eta},$$

and hence

$$\phi = \frac{\bar{\rho}'' - \bar{\lambda}\bar{\rho}'^2}{\bar{\eta}\bar{\rho}'^3}, \tag{6.39}$$

which is identical with (6.28) and is therefore equivalent to (6.6). Equation (6.39), or (6.6), must be satisfied if the C_n in question permits a solution ρ of (6.1) such that $\xi = 0$. But on the other hand, if (6.39) is satisfied, equation (6.1), which is now $\rho_{,ij} + \phi L_{ij} = 0$ by (6.33), is a consequence of (6.38). Therefore we may conclude that for a C_n satisfying (6.4), the solution of (6.1) is identical with the solution $\rho = \rho(\bar{\rho})$ of equation (6.6). Hence Theorem 6.1 is completely established.

7. Applications

In §8 canonical forms for the fundamental tensors of the C_n considered in the preceding section will be obtained. But in the meantime, we apply the preceding results to a $C_n(a)$ which permits a solution of equation (3.1) or (4.6). For convenience we denote by $C_n(a)$ a C_n of constant curvature a. We first show that

THEOREM 7.1. *For a* C_n *whose tensor* L_{ij} *is of the form* (6.3) *or* (6.4), *equation* $(3.1)_1$ *admits infinitely many one-parameter family of solutions* $\rho(x^k, x^m)$ *such that* $\rho(x^k, x_o^m) = 1$ *for some value* x_o^m *of the parameter* x^m.

PROOF. By Theorem 6.1, in the case under consideration the solution of $(3.1)_1$ is identical with the solution $\rho = \rho(\bar{\rho})$ of

$$\rho'' + \bar{\lambda}\rho' + \frac{\bar{\eta}}{n-2}(\rho^{n-1} - \rho) = 0, \tag{7.1}$$

where $\bar{\lambda}$ is a certain function of $\bar{\rho}$. According to the existence theorem, the solution of (7.1) can be written as

$$\rho = (\rho)_o + (\rho')_o(\bar{\rho} - \bar{\rho}_o) + \tfrac{1}{2}(\rho'')_o(\bar{\rho} - \bar{\rho}_o)^2 + \frac{1}{3!}(\rho''')_o(\bar{\rho} - \bar{\rho}_o)^3 + \cdots$$

where $\bar{\rho}_o$ is some value of $\bar{\rho}$; $(\rho)_o$ and $(\rho')_o$ are arbitrary constants; and $(\rho'')_o$, $(\rho''')_o$, etc. are determined from (7.1). If we take $(\rho)_o = 1$ and $(\rho')_o = 0$, then it is readily seen that $(\rho'')_o$, $(\rho''')_o$, etc. are all zero, and consequently the

corresponding solution is $\rho = 1$. Hence, if $f(x^m)$ and $h(x^m)$ are any two functions of x^m, the one-parameter family of solutions $\rho(x^k, x^m)$ of (7.1) corresponding to $(\rho)_o = f(x^m) - f(x_o^m) + 1$ and $(\rho')_o = h(x^m) - h(x_o^m)$ is such that $\rho(x^k, x_o^m) = 1$. Thus our theorem is proved.

As a direct consequence of Theorem 6.1, we have that[15]

THEOREM 7.2. *For a $C_n(a)$, not an S_n, equation* (4.6), *i.e.*

$$\rho_{,ij} + \rho L_{ij} = -\tfrac{1}{2}na\rho g_{ij}, \tag{7.2}$$

admits a solution satisfying $g^{ij}\rho_{,i}\rho_{,j} = 0$, if and only if the tensor L_{ij} of $C_n(a)$ is of the form (6.4).

Also by Theorem 6.1, if, for a $C_n(a)$, not an S_n, equation (7.2) admits a solution satisfying $g^{ij}\rho_{,i}\rho_{,j} \neq 0$, then the tensor L_{ij} of $C_n(a)$ must be of the form (6.3). This condition is also sufficient; indeed, we shall now prove

THEOREM 7.3. *For a $C_n(a)$, not an S_n, equation* (7.2) *admits a solution satisfying $g^{ij}\rho_{,i}\rho_{,j} \neq 0$, if and only if the tensor L_{ij} of $C_n(a)$ is of the form* (6.3). *If $C_n(a)$ satisfies this condition, the solution of* (7.2) *is*

$$\rho = A \exp.\left[-\int \frac{(\bar{\xi} + \tfrac{1}{2}na)\bar{\eta}'}{\bar{\xi}'}\, d\bar{\rho}\right], \qquad A = \text{const.}, \tag{7.3}$$

where the prime denotes differentiation with respect to $\bar{\rho}$.

To prove this theorem we need the formula

$$\bar{\chi}' - \bar{\lambda}\bar{\chi} + \bar{\xi} + \tfrac{1}{2}na = 0, \qquad \text{where} \qquad \bar{\chi} = \frac{\bar{\xi}'}{\bar{\eta}}, \tag{7.4}$$

and $\bar{\lambda} = \bar{\lambda}(\bar{\rho})$ is the function appearing in (6.5), namely,

$$\bar{\rho}_{,ij} = \bar{\lambda}\bar{\rho}_{,i}\bar{\rho}_{,j} + \bar{\chi}g_{ij}. \tag{7.5}$$

Equation (7.4) can be verified as follows. The integrability condition of (7.5) is

$$R^l_{ijk}\bar{\rho}_{,l} = \bar{\lambda}\bar{\rho}_{,i[k}\bar{\rho}_{,j]} + g_{i[j}\bar{\chi}_{,k]}.$$

In consequence of $(1.4)_1$, (6.3) and (7.6), the two members of the above equation reduce to

$$\frac{1}{n-2}(L_{i[j}\bar{\rho}_{,k]} + g_{i[j}L^l_{k]}\bar{\rho}_{,l}) = \frac{1}{n-2}(\bar{\xi}\bar{\eta} + 2\bar{\zeta})g_{i[j}\bar{\rho}_{,k]}$$
$$= -(\bar{\xi} + \tfrac{1}{2}na)g_{i[j}\bar{\rho}_{,k]},$$
$$\bar{\lambda}\bar{\chi}g_{i[k}\bar{\rho}_{,j]} + \bar{\chi}'g_{i[j}\bar{\rho}_{,k]} = (\bar{\chi}' - \bar{\lambda}\bar{\chi})g_{i[j}\bar{\rho}_{,k]},$$

respectively. Equality of these two expressions gives (7.4).

We may now prove our theorem. From the argument below (6.29) it follows that for a $C_n(a)$ satisfying (6.3), equation (7.2) is equivalent to the two equa-

[15] Equation (7.2), for the case of a $C_3(O)$, has been discussed by Chou (**3**, **4**).

tions obtained by putting $\phi = \rho$, $\omega = -\frac{1}{2}na\rho$ in (6.29) and (6.6). These two equations are

$$-\rho' = \frac{(\bar{\xi} + \frac{1}{2}na)\rho}{\bar{\chi}}, \tag{7.6}$$

$$\rho'' + \bar{\lambda}\rho' + \bar{\eta}\rho = 0. \tag{7.7}$$

Differentiating (7.6) and making use of (7.4) we have

$$\begin{aligned} -\rho'' &= \frac{(\bar{\xi} + \frac{1}{2}na)\rho'}{\bar{\chi}} + \frac{\bar{\xi}'\rho}{\bar{\chi}} - \frac{(\bar{\xi} + \frac{1}{2}na)\rho\bar{\chi}'}{\bar{\chi}^2} \\ &= -\frac{(\bar{\xi} + \frac{1}{2}na)^2\rho}{\bar{\chi}^2} + \bar{\eta}\rho - \frac{(\bar{\xi} + \frac{1}{2}na)\rho\bar{\lambda}}{\bar{\chi}} + \frac{(\bar{\xi} + \frac{1}{2}na)^2\rho}{\bar{\chi}^2} \\ &= \bar{\lambda}\rho' + \bar{\eta}\rho. \end{aligned}$$

Therefore, (7.7) is a consequence of (7.6), and hence the solution of (7.2) exists, being given by (7.6). Our theorem is thus proved.

8.1. **Canonical form for the fundamental forms of some C_n**

Let us now return to the general case and proceed to construct in certain privileged coordinates the fundamental forms of the C_n studied in §6. Consider first the case $\xi \neq 0$. If we put $\phi = 0$ in (6.17), we have $\omega = \frac{1}{2}\xi'$. Thus it follows from Theorem 6.1 for $\phi = 0$ that the tensor L_{ij} of a C_n is of the form (6.3), if and only if the equation

$$\rho_{,ij} = \tfrac{1}{2}\xi' g_{ij} \tag{8.1}$$

admits a solution ρ. Following Brinkmann (**2,** 123–4), we can prove that in any V_n for which equation (8.1) admits a solution ρ, a coordinate system exists in which $x^n = \rho$ and the fundamental form of V_n has the form

$$\xi\, h_{bc}(x^d)\, dx^b\, dx^c + \frac{(d\rho)^2}{\xi}, \qquad b, c, d = 1, \cdots, n-1, \tag{8.2}$$

where $h_{bc}(x^d)$ is independent of ρ. Conversely, for a V_n with fundamental form (8.2), where ξ is any function of ρ, ρ satisfies (8.1).

We first suppose that $\xi' \neq 0$. Put $z^{-2} = e\xi > 0$, where $e = \pm 1$, and let e be absorbed in $h_{bc}(x^d)$. Then the form (8.2) becomes

$$z^{-2}\left[h_{bc}(x^d)\, dx^b\, dx^c + \frac{(dz)^2}{Z}\right], \tag{8.3}$$

where we have written $Z = Z(z) = ez^{-4}\left(\frac{d\rho}{dz}\right)^{-2}$. Now *the condition for* (8.3) *to represent a C_n is that the form $h_{bc}\, dx^b\, dx^c$ represents an S_{n-1}.* In fact, (8.3) is conformal to the separable form

$$'g_{ij}\, dx^i\, dx^j = h_{bc}(x^d)\, dx^b\, dx^c + \frac{(dz)^2}{Z}, \tag{8.4}$$

and it was shown by Ficken (**9,** 897) that the condition for (8.4) to represent a C_n is equivalent to the condition for $h_{bc}\,dx^b\,dx^c$ to represent an S_{n-1}.

We now calculate the scalar curvature of (8.3). It is known that the scalar curvatures C and $'C$ of (8.3) and (8.4) are connected by (cf. 2.13))

$$C = {}'Cz^2 + \frac{2}{n}\,z\,{}'g^{ij}z_{/ij} - {}'g^{ij}z_{/i}z_{/j},$$

where the solidus denotes covariant differentiation with respect to (8.4). But $z_{/b} = 0$, $z_{/n} = 1$, and, if ${}'\Gamma^n_{nn}$ is a component of the Christoffel symbol of the second kind for (8.4),

$${}'g^{ij}z_{/ij} = {}'g^{nn}z_{/nn} = Z(-{}'\Gamma^n_{nn}) = \tfrac{1}{2}Z'.$$

Therefore we have

$$C = {}'Cz^2 + \frac{1}{n}\,zZ' - Z. \tag{8.5}$$

To express $'C$ in terms of the scalar curvature B of h_{bc}, we note that the non-vanishing components of the Ricci tensor of (8.4) are ${}'R_{bc} = \overset{(h)}{R}_{bc}$, where the superindex (h) indicates that the tensor is refered to h_{bc}. Therefore it follows that $n'C = (n-2)B$. Using this in (8.5) we have

$$nC = (n-2)Bz^2 + zZ' - nZ = (n-2)Bz^2 + z^{n+1}(z^{-n}Z)'. \tag{8.6}$$

If h_{bc} represents an S_{n-1}, B = const. and (8.6) shows that $C = C(z)$. Then, by integrating (8.6) we have

$$Z = Az^n + Bz^2 + nz^n \int Cz^{-(n+1)}\,dz, \tag{8.7}$$

where A is a constant.

We now suppose that $\xi' = 0$. Then since by hypothesis ξ = const. $\neq 0$, an obvious transformation would reduce the fundamental form (8.2) to the form

$$h_{bc}(x^d)\,dx^b\,dx^c + e\,(dz)^2, \qquad e = \pm 1.$$

Like (8.3), this represents a $C_n(C)$ if and only if h_{bc} represents an $S_{n-1}(B)$; and then $nC = (n-2)B$ = const.

Summing up the preceding results and using the Riemann form for the fundamental form of an S_{n-1}, we have

THEOREM 8.1. *A V_n is a C_n whose tensor L_{ij} is of the form* (6.3) *if and only if its fundamental form can be reduced to either of the following forms*:

$$\frac{\sum_a e_a(dx^a)^2}{z^2\left[1 + \dfrac{B}{4}\sum_a e_a(x^a)^2\right]^2} + \frac{(dz)^2}{z^2\left[Az^n + Bz^2 + nz^n\displaystyle\int Cz^{-(n+1)}\,dz\right]}, \tag{8.8}$$

$$\frac{\sum_a e_a(dx^a)^2}{\left[1+\frac{B}{4}\sum_a e_a(x^a)^2\right]^2}+e_n(dz)^2, \tag{8.9}$$

where $a = 1, \cdots, n-1$; *each* e *is* ± 1; A *and* B *are constants; and* $C = C(z)$. *The scalar curvature of* C_n *is* C *for* (8.8) *and is* $\frac{n-2}{n}B$ *for* (8.9).

In particular, when $C = \text{const.}$, we have

COROLLARY 8.1. *A* V_n *is a* C_n *whose tensor* L_{ij} *is of the form* (6.3) *and whose scalar curvature* C *is constant, if and only if its fundamental form can be reduced to either of the following forms*:

$$\frac{\sum_a e_a(dx^a)^2}{z^2\left[1+\frac{B}{4}\sum_a e_a(x^a)^2\right]^2}+\frac{(dz)^2}{z^2(Az^n+Bz^2-C)} \tag{8.8'}$$

$$\frac{\sum_a e_a(dx^a)^2}{\left[1+\frac{nC}{4(n-2)}\sum_a e_a(x^a)^2\right]^2}+e_n(dz)^2. \tag{8.9'}$$

We remark that *in* (8.8) *or* (8.8'), *if the constant* A *is not zero, it may be supposed to be* ± 1. In fact, if we put $z = \alpha' z$, $\alpha = \text{const.}$, (8.3) and (8.7) become

$$'z^{-2}\left[\alpha^{-2}h_{bc}(x^d)\,dx^b\,dx^c+\frac{(d'z)^2}{'Z}\right],$$

$$'Z = A\alpha^n\,'z^n+B\alpha^2\,'z^2+n\,'z^n\int C\,'z^{-(n+1)}\,d'z.$$

Now if $A \neq 0$, α can be so chosen that $A\alpha^n = \pm 1$. But since the scalar curvature of $(\alpha^{-2}h_{bc})$ is equal to α^2 times that of h_{bc}, this transformation from z to $'z$ does not affect the property of h_{bc} being the fundamental tensor of an $S_{n-1}(B)$. Hence our assertion is proved.

8.2. Continuation

We now consider the case $\xi = 0$. By Theorem 6.1, the tensor L_{ij} of a $C_n(0)$ is of the form (6.4) if and only if the equations

$$\rho_{,ij} = 0, \qquad \xi = g^{ij}\rho_{,i}\rho_{,j} = 0 \tag{8.10}$$

admit a solution for ρ. Then ρ is a function of $\bar{\rho}$ and we have

$$L_{ij} = \eta(\rho)\rho_{,i}\rho_{,j}, \qquad \eta \neq 0. \tag{8.11}$$

Brinkmann (**2,** 131–2) proved that in consequence of (8.10) a coordinate system x^p, y, z ($p, q, r, s = 1, \cdots, n-2$) exists in which $\rho = z$ and the fundamental form of $C_n(0)$ has the form

$$g_{ij}\,dx^i\,dx^j = f_{pq}\,dx^p\,dx^q+2m_p\,dx^p dz+m(dz)^2+2\,dy\,dz, \tag{8.12}$$

where the functions f_{pq}, m_p and m are independent of y. For (8.12), the non-vanishing components of Γ^k_{ij} are

$$\Gamma^r_{pq} = \overset{(f)}{\Gamma}{}^r_{pq}, \qquad \Gamma^r_{pn}, \qquad \Gamma^r_{nn}, \qquad \Gamma^{n-1}_{pq}, \qquad \Gamma^{n-1}_{nn}, \qquad \Gamma^{n-1}_{pn-1}, \qquad \Gamma^{n-1}_{pn},$$

where the superindex (f) indicates that the quantities are referred to the form $f_{pq}\,dx^p\,dx^q$, z being considered as a parameter. From these it follows that

$$R^s_{pqr} = \partial_{[q}\Gamma^s_{r]p} - \Gamma^i_{p[q}\Gamma^s_{r]i} = \overset{(f)}{R}{}^s_{pqr}. \tag{8.13}$$

On the other hand, since $C_n(0)$ is conformal-Euclidean, we have from $(1.4)_1$ and (8.11) that

$$R^l_{ijk} = Z(z_{,i}z_{,[j}\delta^l_{k]} + g_{i[j}z_{,k]}z_{,}{}^l), \tag{8.14}$$

where $z_{,}{}^l = g^{li}z_{,i}$, and Z is defined by $\eta(\rho) \equiv (n-2)Z(z)$. But

$$z_{,p} = z_{,n-1} = 0, \qquad z_{,n} = 1; \qquad z_{,}{}^p = z_{,}{}^n = 0, \qquad z_{,}{}^{n-1} = 1. \tag{8.15}$$

Therefore the components R^s_{pqr} of R^l_{ijk} are zero. Thus (8.13) gives $\overset{(f)}{R}{}^s_{pqr} = 0$, and consequently, for each value of z the form $f_{pq}\,dx^p\,dx^q$ in (8.12) represents a Euclidean V_{n-2}. Hence by means of a suitable transformation of the form

$$'x^p = {}'x^p(x^r, z), \qquad 'y = y, \qquad 'z = z,$$

the fundamental form (8.12) can be reduced to (after dropping the dash)

$$g_{ij}\,dx^i\,dx^j = \sum_p e_p(dx^p)^2 + 2m_p\,dx^p\,dz + m(dz)^2 + 2\,dy\,dz, \quad e^p = \pm 1, \tag{8.16}$$

where, as in (8.12), the functions m_p and m are independent of y.

For this fundamental form we have

$$\Gamma^s_{pq} = \Gamma^s_{qn-1} = \Gamma^s_{nn-1} = \Gamma^n_{pq} = \Gamma^n_{pn} = 0, \qquad \Gamma^s_{pn} = \tfrac{1}{2}e_s\partial_{[p}m_{s]},$$

and whence $R^s_{pqn} = \partial_{[q}\Gamma^s_{n]p} - \Gamma^i_{p[q}\Gamma^s_{n]i} = \partial_q\Gamma^s_{np} = \tfrac{1}{2}e_s\partial_q(\partial_{[p}m_{s]})$. But R^s_{pqn} should be zero by (8.14) and (8.15). Therefore we have

$$\partial_{[p}m_{s]} = P_{ps}(z) = -P_{sp}(z), \tag{8.17}$$

where $P_{ps}(z)$ are some functions of z alone.

It is easily seen that if $Y(x^r, z)$ is a function of x^r and z, and $Q^p_r(z)$ are functions of z satisfying the equations

$$\sum_r e_r Q^p_r Q^q_r = e_p\delta_{pq}, \tag{8.18}$$

or their equivalent

$$\sum_p e_p Q^p_r Q^p_q = e_r\delta_{rq}, \tag{8.18'}$$

then the following change of coordinates

$$'x^p = Q^p_r(z)x^r, \qquad 'y = y + Y(x^r, z), \qquad 'z = z \tag{8.19}$$

will preserve the form of (8.16). Now for (8.16) the contravariant components g^{ij} of g_{ij} are

$$
\begin{aligned}
&g^{np} = 0, \qquad g^{nn-1} = 1, \qquad g^{nn} = 0; \\
&g^{n-1n-1} = g^{n-1n-1}, \qquad g^{pn-1} = -e_p m_p, \qquad g^{pq} = e_p \delta_{pq}.
\end{aligned} \tag{8.20}
$$

Therefore if we denote by $'g^{pn-1}$ and $'m_p$ the corresponding values to g^{pn-1} and m_p in this new coordinate system (8.19), we have

$$
\begin{aligned}
-e_p\, 'm_p = {}'g^{pn-1} &= g^{ij} \frac{\partial}{\partial x^i}(Q_r^p x^r) \frac{\partial}{\partial x^j}(y + Y) \\
&= \sum_r e_r Q_r^p \frac{\partial Y}{\partial x^r} - \sum_r e_r m_r Q_r^p + (Q_r^p)' x^r.
\end{aligned} \tag{8.21}
$$

We shall now prove that the functions Y and Q_r^p may be so chosen that $'m_p = 0$.

Making use of the conjugate system $\bar{Q}_p^q$ of Q_r^p defined by

$$
Q_r^p \bar{Q}_p^q = \delta_r^q, \tag{8.22}
$$

it follows from (8.21) that the condition $'m_p = 0$ is equivalent to

$$
\frac{\partial Y}{\partial x^q} - m_q + e_q \bar{Q}_p^q (Q_r^p)' x^r = 0, \qquad (q \text{ not summed}).
$$

But we have from (8.18) and (8.22) that $e_p \bar{Q}_p^q = e_q Q_q^p$, therefore the above equations can be written

$$
\frac{\partial Y}{\partial x^q} = m_q - \sum_p e_p Q_q^p (Q_r^p)' x^r. \tag{8.23}
$$

Now let us consider the variable z in Y as a parameter. Then the integrability condition of (8.23) is

$$
\partial_{[r} m_{q]} = \sum_p e_p Q_{[q}^p (Q_{r]}^p)',
$$

which becomes, by (8.17),

$$
\sum_p e_p Q_{[q}^p (Q_{r]}^p)' = P_{rq}. \tag{8.24}
$$

If we write $\bar{n} = n - 2$, then (8.24) are $\frac{1}{2}\bar{n}(\bar{n} - 1)$ equations, which, together with the $\frac{1}{2}\bar{n}(\bar{n} + 1)$ equations (8.18), form a system of $\bar{n}^2$ equations for the $\bar{n}^2$ unknowns Q_r^p. That this system of equations is compatible we shall prove later. For the moment, we assume that a set of values Q_r^p have been found satisfying (8.18) and (8.24). Then, equations (8.23) can be solved for Y. Hence the functions Q_r^p and Y in (8.19) can be so determined that $'m_p = 0$, as was to be proved. In other words, there exists a coordinate system in which the fundamental form of $C_n(0)$ has the form

$$
g_{ij}\, dx^i\, dx^j = \sum_p e_p (dx^p)^2 + m(dz)^2 + 2\, dy\, dz. \tag{8.25}
$$

Now we have to determine m so that (8.14) is satisfied. For (8.25) the non-vanishing components of Γ^k_{ij} and R^l_{ijk} are[16]

$$\Gamma^p_{nn} = -\tfrac{1}{2}e_p\partial_p m, \qquad \Gamma^{n-1}_{nn} = -\tfrac{1}{2}\partial_n m, \qquad \Gamma^{n-1}_{pn} = \tfrac{1}{2}\partial_p m;$$
$$R^s_{nnr} = \tfrac{1}{2}e_s\partial_r\partial_s m, \qquad R^{n-1}_{pqn} = \tfrac{1}{2}\partial_q\partial_p m.$$

On the other hand, we have from (8.14) and (8.15) that the only non-vanishing components of R^l_{ijk} are

$$R^s_{nnr} = Z\delta^s_r, \qquad R^{n-1}_{pqn} = Zg_{pq}.$$

Comparison of these two different expressions for the components of R^l_{ijk} shows that

$$\partial_q\partial_p m = 2Zg_{pq} = 2Ze_p\delta_{pq},$$

which gives, since m is independent of y,

$$m = Z\sum_p e_p(x^p)^2 + Z_p x^p + Z_{n-1}, \tag{8.26}$$

where the Z's are some functions of z alone.

We can prove, by direct computations, that the fundamental form (8.25) with the value (8.26) for m is actually the fundamental form of a $C_n(0)$ whose tensor L_{ij} is of the form (6.4). Indeed, for this fundamental form the only non-vanishing components of R^l_{ijk} and R_{ij} are

$$R^s_{nnr} = Z\delta^s_r, \qquad R^{n-1}_{pqn} = Zg_{pq}; \qquad R_{nn} = (n-2)Z.$$

From these it follows that $L = R = 0$, $L_{ij} = R_{ij}$, and consequently, we can easily verify that conditions (1.4) are satisfied.

Hence we may state the following

THEOREM 8.2. *A V_n is a $C_n(0)$ whose tensor L_{ij} is of the form* (6.4) *if and only if its fundamental form can be reduced to the form*

$$\sum_p e_p(dx^p)^2 + 2\,dy\,dz + [Z\sum_p e_p(x^p)^2 + Z_p x^p + Z_{n-1}](dz)^2, \tag{8.27}$$
$$e_p = \pm 1, \qquad p = 1, \cdots, n-2,$$

where the Z's are some functions of z alone.

To complete the proof of this theorem, we now show, as was asserted, that equations (8.18) and (8.24) can actually be solved for Q^p_r. From the equivalent (8.18′) of (8.18) we have

$$(\sum_p e_p Q^p_r Q^p_q)' = \sum_p e_p(Q^p_r)'(Q^p_q) + \sum_p e_p Q^p_r(Q^p_q)' = 0. \tag{8.28}$$

Hence (8.24) can be written

$$2\sum_p e_p(Q^p_q)(Q^p_r)' = P_{rq}.$$

Transvecting these by $\bar{Q}^q_s$, we find

$$2e_s(Q^s_r)' = P_{rq}\bar{Q}^q_s \qquad (s \text{ not summed}),$$

that is, $$(Q^s_r)' = \tfrac{1}{2}e_s\bar{Q}^q_s P_{rq} = \tfrac{1}{2}e_s\sum_q e_p e_s Q^s_q P_{rq} \quad (s \text{ not summed}).$$

[16] Strictly speaking, there are two more non-vanishing components of R^l_{ijk}, namely, R^s_{nrn} $(= -R^s_{nnr})$ and R^{n-1}_{pnq} $(= -R^{n-1}_{pqn})$.

Therefore, on account of (8.18), equations (8.24) are equivalent to

$$(Q_r^s)' = \tfrac{1}{2}\sum_q e_q Q_q^s P_{rq} . \tag{8.29}$$

These are a system of $\bar{n}^2 = (n-2)^2$ ordinary linear differential equations in the $\bar{n}^2$ unknowns Q_r^s and one independent variable z. They have a unique solution for any set of initial values $(Q_r^s)_o$ at $z = z_o$. We can easily prove that when $(Q_r^s)_o$ are suitably chosen, the solution Q_r^s satisfies (8.18). Indeed, since $(Q_r^s)_o$ can always be found satisfying (8.18), we need only show that the expression (8.28) is satisfied because of (8.29). Now

$$\begin{aligned}(\sum_p e_p Q_r^p Q_q^p)' &= \sum_p e_p (Q_q^p)(\tfrac{1}{2}\sum_s e_s Q_s^p P_{rs}) + \sum_p e_p (Q_r^p)(\tfrac{1}{2}\sum_s e_s Q_s^p P_{qs}) \\ &= \tfrac{1}{2}(\sum_s e_q \delta_{qs} e_s P_{rs} + \sum_s e_r \delta_{rs} e_s P_{qs}) \\ &= \tfrac{1}{2}(P_{rq} + P_{qr}),\end{aligned}$$

which is zero by (8.17). Hence our assertion is proved.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

REFERENCES

1. BRINKMANN, H. W., "*Riemann spaces conformal to Einstein spaces*," Math. Ann., 91 (1924), 269–278.
2. BRINKMANN, H. W., "*Einstein spaces which are mapped conformally on each other*," Math. Ann., 94 (1925), 119–145.
3. CHOU, P. Y., "*Isotropic static solutions of the field equation in Einstein's theory of gravitation*," American J. Math., 59 (1937), 754–763.
4. CHOU, P. Y., "*On the method of finding isotropic static solutions of Einstein's field equations of gravitation*," American J. Math., 62 (1940), 43–48.
5. EISENHART, L. P., Riemannian geometry (1926), Princeton.
6. EISENHART, L. P., Continuous groups of transformations (1934), Princeton.
7. FIALKOW, A., "*Totally geodesic Einstein spaces*," Bull. American Math. Soc., 45 (1939), 423–428.

7′. FIALKOW, A., "*Correction to 'Totally geodesic Einstein spaces,'*" loc. cit., 48 (1942), 167–168.

8. FIALKOW, A., "*Conformal geodesics*," Trans. American Math. Soc., 45 (1939), 443–473.
9. FICKEN, F. A., "*The Riemannian and affine differential geometry of product-spaces*," Ann. Math., 40 (1939), 892–913.
10. SCHOUTEN, J. A., AND STRUIK, D. J., "*On some sroperties of general manifolds relating to Einstein's theory of graviation*," American J. Math., 43 (1921), 217–221.
11. SCHOUTEN, J. A., AND STRUIK, D. J., Einführung in die neueren Methoden der Differentialgeometrie, II, (1938), Groningen.
12. YANO, K., "*Concircular geometry, I. Concircular transformations*," Proc. Imp. Acad. Tokyo, 16 (1940), 195–200.
13. YANO, K., "*Concircular geometry, II. Integrability conditions of* $\rho_{\mu\nu} = \phi g_{\mu\nu}$," loc. cit., 354–360.
14. WONG, YUNG-CHOW, "*Some Einstein spaces with conformally separable fundamental tensors*," in course of publication in the Transactions of the American Mathematical Society.

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

EUCLIDEAN APPLICATIONS OF THE PROJECTIVE DIFFERENTIAL GEOMETRY OF THE R_λ-CORRESPONDENT

BY P. O. BELL AND W. C. FOREMAN

(Received July 3, 1942)

INTRODUCTION

Let C_λ denote an arbitrary curve on a surface S in ordinary projective space. P. O. Bell has recently presented the following construction for the line which he calls the *R_λ-correspondent*[1] of the tangent to C_λ at a point x: Let X denote a point of C_λ distinct from x, let U^1, U^2 denote, respectively, the points of intersection of the asymptotic u^1- and u^2-curves passing through x with the asymptotic u^2- and u^1-curves passing through X, and let W denote the point of intersection of the tangent plane to S at x with the line joining the points U^1, U^2. *The R_λ-correspondent of the tangent to C_λ at x is the limit of the line joining x and W as X approaches x along C_λ.* He has proved, moreover, the following three theorems:

(i) A curve C_λ is a curve of Darboux if and only if at each of its points the R_λ-correspondent of the tangent to C_λ coincides with this tangent;

(ii) A curve C_λ is a curve of Segre if and only if at each of its points the tangent to C_λ and its R_λ-correspondent are conjugate tangents of S;

(iii) A curve C_λ is a curve of Segre if and only if for a general point x of C_λ the limit of W as X tends to x along C_λ is a point W_0 distinct from x. The point W_0 is the intersection of the directrix of the first kind of Wilczynski with the tangent at x to the corresponding curve C_λ of Darboux.

The correspondent whose definition results from that of the R_λ-correspondent upon the removal of the restriction that the asymptotic net be parametric will hereafter be called the R_λ-correspondent. This paper is concerned largely with the study of this correspondent and of a dual correspondent which will be called the R_λ^*-correspondent. In chapter I the study is from the point of view of projective differential geometry. The determination of the R_λ- correspondent for a general parametric net leads naturally to projective definitions of a number of interesting nets as well as of a rectilinear congruence covariantly determined with respect to the parametric net. Among the nets thus determined are the R_λ-net, the principal-associate net of the parametric net and the associate conjugate net of the parametric net. In Chapter II euclidean geometric characterizations of the R_λ- and R_λ^*-correspondents are first obtained. These characterizations form the basis for the formulation of dual theories in euclidean differential geometry of nets. A unique parametric net is discovered which is such that every associated R_λ-net is orthogonal. This net consists of curves

[1] P. O. Bell, Bull. Am. Math. Soc. 47, 1941, pp. 509–511. For a detailed study of the R_λ-correspondent and of associated R_λ-derived curves for the case in which the asymptotic curves are parametric see Trans. Am. Math. Soc. 46.

whose normal curvatures at a point are equal to the mean of the principal normal curvatures at the point. We call these curves the lines of mean curvature of S. The dual of this net is the unique conjugate parametric net which is such that every associated R_λ^*-net is conjugate. This net is the mean conjugate net and consists of those curves of S whose radii of normal curvature at a point x are equal to the mean of the principal radii of normal curvature at the point. Finally euclidean characterizations are obtained for the curves of Segre and the curves of Darboux and simple new determinations are given for the lines of curvature.

I. Projective Differential Geometry

1. The R_λ-correspondent

Let the homogeneous coordinates of a general point x of an analytic surface S and of a corresponding point z not in the tangent plane to S at x be functions x^i and z^i $(i = 1, 2, 3, 4)$, of two independent variables u^1, u^2. The pairs of functions (x, z) are solutions of a system of differential equations of the form[2]

$$\begin{aligned} \frac{\partial^2 x}{\partial u^i \partial u^j} &= \binom{k}{ij} \frac{\partial x}{\partial u^k} + a_{ij} x + d_{ij} z, \\ \frac{\partial z}{\partial u^i} &= b_i^k \frac{\partial x}{\partial u^k} + c_i x + e_i z, \qquad (i, j, k = 1, 2), \end{aligned} \tag{1.1}$$

whose coefficients are functions of u^1, u^2 which satisfy certain integrability conditions. The usual summation convention of tensor analysis is employed throughout this paper.

The asymptotic curves are the integral curves of the differential equation

$$d_{ij} du^i du^j = 0. \tag{1.2}$$

An arbitrary one parameter family F_λ of curves is defined by the curvilinear equation

$$du^2 - \lambda du^1 = 0, \tag{1.3}$$

where λ is an arbitrary function of u^1, u^2. Let C_λ denote the curve of the family which passes through the point x. The curvilinear coordinates of a general point on a curve of S may be represented by the functions $u^i = u^i(t)$, $(i = 1, 2)$, of a parameter t. If this curve passes through x it will be the curve C_λ if for a general value of t

$$\lambda = \frac{du^2}{dt} \Big/ \frac{du^1}{dt}. \tag{1.4}$$

Let us determine the direction of the R_λ-correspondent as defined in the Introduction for a general non-asymptotic parametric net. The curvilinear coordinates of the point X are given by $u^1(t + \Delta t)$, $u^2(t + \Delta t)$. The points

[2] Cf. E. P. Lane, *Projective Differential Geometry of Curves and Surfaces* (hereafter referred to as P. D. G.), University of Chicago Press, 1932, p. 183, equations (35).

U^1 and U^2 are therefore given by $(u^1(t + \Delta t), u^2)$ and $(u^1, u^2(t + \Delta t))$. The general homogeneous coordinates of the points U^i are consequently functions of t, and may be represented by the development

$$U^i = x + \frac{\partial x}{\partial u^i}\frac{du^i}{dt}\Delta t + \left(\frac{\partial^2 x}{\partial u^{i^2}}\left(\frac{du^i}{dt}\right)^2 + \frac{\partial x}{\partial u^i}\frac{d^2u^i}{dt^2}\right)(\Delta t)^2/2$$

$$+ \left(\frac{\partial^3 x}{\partial u^{i^3}}\left(\frac{du^i}{dt}\right)^3 + 3\frac{\partial^2 x}{\partial u^{i^2}}\frac{du^i}{dt}\frac{d^2u^i}{dt^2} + \frac{\partial x}{\partial u^i}\frac{d^3u^i}{dt^3}\right)(\Delta t)^3/6 \tag{1.5}$$

$$+ \left(\frac{\partial^4 x}{\partial u^{i^4}}\left(\frac{du^i}{dt}\right)^4 + 6\frac{\partial^3 x}{\partial u^{i^3}}\left(\frac{du^i}{dt}\right)^2\frac{d^2u^i}{dt^2} + f_1^i\frac{\partial^2 x}{\partial u^{i^2}} + f_2^i\frac{\partial x}{\partial u^i}\right)(\Delta t)^4/24 + \cdots,$$

wherein f_j^i, $(i, j = 1, 2)$, represent functions of u^1, u^2 which for our purposes do not require explicit determination, and where i is not summed.

By differentiating equations (1.1) we find that the coefficient of z in the expression for $\frac{\partial^3 x}{\partial u^{i^3}}$ is $\frac{\partial d_{ii}}{\partial u^i} + d_{ii}e_i + d_{ij}(_{ii}^{j})$. The coefficient of z in the expression for the homogeneous coordinates of the point U^i is, therefore $d_{ii}\left(\frac{du^i}{dt}\right)^2(\Delta t)^2/2 + \omega_i\left(\frac{du^i}{dt}\Delta t\right)^3/6 + \cdots$, wherein $\omega_i = \frac{\partial d_{ii}}{\partial u^i} + d_{ii}e_i + d_{ij}(_{ii}^{j}) + 3d_{ii}\left(\frac{d^2u^i}{dt^2}\right)\Big/\left(\frac{du^i}{dt}\right)^2$. The point W has homogeneous coordinates which may be obtained by forming a linear combination of those of U^1 and U^2 which contains no z term. Such a combination is

$$\left(3d_{22}\left(\frac{du^2}{dt}\right)^2 + \omega_2\left(\frac{du^2}{dt}\right)^3\Delta t\right)U^1 - \left(3d_{11}\left(\frac{du^1}{dt}\right)^2 + \omega_1\left(\frac{du^1}{dt}\right)^3\Delta t\right)U^2.$$

Expanding this, we obtain the expression

$$3\left(d_{22}\left(\frac{du^2}{dt}\right)^2 - d_{11}\left(\frac{du^1}{dt}\right)^2\right)x + \left(\omega_2\left(\frac{du^2}{dt}\right)^3 - \omega_1\left(\frac{du^1}{dt}\right)^3\right)(\Delta t)x$$

$$+ 3(\Delta t)\left(d_{22}\left(\frac{du^2}{dt}\right)^2\left(\frac{du^1}{dt}\right)\frac{\partial x}{\partial u^1} - d_{11}\left(\frac{du^1}{dt}\right)^2\left(\frac{du^2}{dt}\right)\frac{\partial x}{\partial u^2}\right) \tag{1.6}$$

$$+ \text{terms of order } (\Delta t)^2$$

for the homogeneous coordinates of the point W. Therefore, if the parametric curves are not asymptotic curves, the direction of the R_λ-correspondent of the tangent to C_λ at x is defined by

$$du^2/du^1 = -d_{11}/d_{22}\lambda. \tag{1.7}$$

Let C_W denote the curve described by the point W as Δt varies. The expression (1.6) for the coordinates of W shows that *if the parametric net is not the asymptotic net and the curve C_λ is not an integral curve of the equation*

$$d_{11}(du^1)^2 - d_{22}(du^2)^2 = 0 \tag{1.8}$$

[3] P. O. Bell, loc. cit.

then the R_λ-correspondent of the tangent to C_λ at x is tangent at x to the corresponding curve C_W, that is to say, the point x is the limit of W as Δt tends to zero.

If C_λ is an integral curve of (1.8), the limit of W as Δt tends to zero is a point W_0, distinct from x, whose homogeneous coordinates are given by

$$(1.9)\quad \left(\omega_2\left(\frac{du^2}{dt}\right)^3 - \omega_1\left(\frac{du^1}{dt}\right)^3\right)x + 3\left(d_{22}\left(\frac{du^2}{dt}\right)^2\frac{du^1}{dt}\frac{\partial x}{\partial u^1} - d_{11}\left(\frac{du^1}{dt}\right)^2\frac{du^2}{dt}\frac{\partial x}{\partial u^2}\right),$$

wherein $d_{22}\left(\frac{du^2}{dt}\right)^2 = d_{11}\left(\frac{du^1}{dt}\right)^2$. If we evaluate (1.9) explicitly for the directions $\lambda = \pm(d_{11}/d_{22})^{\frac{1}{4}}$, we obtain, except for unessential factors

$$(1.10)\quad 3\frac{\partial x}{\partial u^1} - \left(\frac{\partial \log (d_{11}d_{22}^3)^{\frac{1}{4}}}{\partial u^1} + e_1 + \frac{d_{1j}(_1{}^j{}_1)}{d_{11}}\right)x \pm \left(\frac{d_{11}}{d_{22}}\right)^{\frac{1}{4}}\left(3\frac{\partial x}{\partial u^2} - \left(\frac{\partial \log (d_{11}^3d_{22})^{\frac{1}{4}}}{\partial u^2} + e_2 + \frac{d_{2j}(_2{}^j{}_2)}{d_{22}}\right)x\right).$$

These two points W_0 determine a line which will be called the w-line associated with the parametric net at x.

For the parametric asymptotic net since $d_{11} = d_{22} = 0$ and $d_{12} \neq 0$, the directions of the R_λ-correspondent of the tangent to C_λ at x is determined by the terms of order $(\Delta t)^2$ in (1.6). This direction, which is defined by

$$(1.11)\quad du^2/du^1 = -\,(_1{}^2{}_1)/(_2{}^1{}_2)\lambda^2$$

has been studied on previous occasions[3] from a projective standpoint.

For an arbitrarily selected parametric net there exists in correspondence with any one parameter family of curves F_λ of S a family of curves $F_{\bar\lambda}$ characterized by the property that at a general point x of S the tangent to the curve $C_{\bar\lambda}$ through x is the R_λ-correspondent of the tangent to C_λ at x. The family $F_{\bar\lambda}$, thus defined in association with a family F_λ and a given parametric net, will be called the *family of R_λ-derived curves*. If the parametric net is not the asymptotic net, equation (1.7), with u^1, u^2 allowed to vary, defines the family of R_λ-derived curves. If, however, the asymptotic net is parametric, equation (1.11) defines the family of R_λ-derived curves. The net which consists of the family F_λ and the family of R_λ-derived curves will be called the R_λ-net.

2. Associate nets of a parametric net

The correspondence defined by (1.7), between the direction of the tangent to C_λ at x and its R_λ-correspondent is an involution. At a point x of S, the double elements of this involution are the solutions of the equation

$$(2.1)\quad d_{ii}(du^i)^2 = 0.$$

The net of integral curves of this equation will be called the *principal associate net* of the parametric net. Since the directions at x of the curves of this net are double elements of the involution (1.7), we have immediately

THEOREM (2.1) *At a general point x of S the tangents of the principal associate net of the parametric net are harmonic conjugates with respect to the tangent at x to an arbitrary curve C_λ through x and the R_λ-correspondent of this tangent.*

If $d_{12} = 0$, the integral curves of (2.1) are the asymptotic curves of S. Hence we have

THEOREM (2.2) *The asymptotic net is the principal associate net of an arbitrary conjugate parametric net.*

The unique conjugate net whose tangents at a general point x separate harmonically the tangents at x to the parametric curves is called the *associate conjugate net* of the parametric net. The differential equation of this net is equation (1.8). Hence we have

THEOREM (2.3) *The associate conjugate net of a parametric net consists of those curves, a general one C_λ of which has the property that, the limit of the point W as X tends toward x along C_λ is a point W_0 distinct from x. The two points W_0 which correspond to the two curves of the net which pass through x determine the w-line of the parametric net at x.*

The necessary and sufficient condition that the tangent to C_λ at x and its R_λ-correspondent be conjugate tangents is that the harmonic invariant of the two forms

$$d_{ij}du^i du^j \quad \text{and} \quad d_{11}\lambda(du^1)^2 + (d_{22}\lambda^2 - d_{11})du^1 du^2 - d_{22}(du^2)^2$$

vanish; that is

$$d_{12}(d_{22}\lambda^2 - d_{11}) = 0. \tag{2.2}$$

If $d_{12} = 0$, equation (2.2) is satisfied independently of λ. Hence we have that

THEOREM (2.4) *Every R_λ-net derived in association with an arbitrary conjugate parametric net is a conjugate net.*

The equation $d_{22}\lambda^2 - d_{11} = 0$ is not only the equation for the directions at x of the curves of the associate conjugate net of the parametric net, but it is the necessary and sufficient condition that the tangents at x of the R_λ-net separate harmonically the tangents to the parametric net. Hence we have

THEOREM (2.5) *Corresponding to a non-asymptotic parametric net there is a unique R_λ-net whose tangents at a general point x separate harmonically the tangents to the parametric curves at the point. This R_λ-net is the associate conjugate net of the parametric net.*

If C_λ is an asymptotic curve of S it is clear from equation (1.2) that the direction of the other asymptotic curve at a general point x of C_λ is given by $\dfrac{du^2}{du^1} = \dfrac{d_{11}}{d_{22}\lambda}$. We have, therefore

THEOREM (2.6) *The tangents at x to the parametric curves of S separate harmonically the R_λ-correspondent of the tangent at x to an asymptotic curve C_λ and the tangent to the other asymptotic curve at x.*

The harmonic conjugate of the tangent at x to an arbitrary curve C_λ of S with respect to the tangents at x to the curves of any selected parametric net

may be geometrically characterized in the following simple manner. Let x and X denote two points on a curve C_λ of S, and let U^1 and U^2 denote the points of intersection of the parametric u^1- and u^2-curves passing through x with the parametric u^2- and u^1-curves passing through X. *The limit of the line joining U^1 and U^2 as X approaches x along C_λ is the harmonic conjugate of the tangent to C_λ at x with respect to the tangents at x to the parametric curves of S.*

The validity of the above construction may be proved as follows. The curvilinear coordinates of the points U^1 and U^2 are $\bar{u}^1$, u^2 and u^1, $\bar{u}^2$, respectively, wherein $\bar{u}^1(t) = u^1(t + \Delta t)$, $\bar{u}^2(t) = u^2(t + \Delta t)$. The coordinates u^1, $\bar{u}^2$ of U^2 may be defined in terms of those of U^1 by the relations $u^1(t) = \bar{u}^1(t - \Delta t)$, $\bar{u}^2(t) = u^2(t + \Delta t)$. Except for terms of order $(\Delta t)^2$ these coordinates are $\bar{u}^1 - \dfrac{du^1}{dt}\,\Delta t$, $u^2 + \dfrac{du^2}{dt}\,\Delta t$. Hence, the direction of the limit of the line joining U^1, U^2 as Δt tends to zero is given by $\dfrac{du^2}{du^1} = -\dfrac{du^2}{dt} \Big/ \dfrac{du^1}{dt} = -\lambda$. In view of the cross ratio equation $(0, \infty, \lambda, -\lambda) = -1$ the proof is complete.

As an important special case of the above construction we have

THEOREM (2.7) *If the asymptotic net is parametric, the limit of the line joining U^1, U^2 as X tends to x along C_λ is the conjugate of the tangent to C_λ at x.*

II. APPLICATIONS TO EUCLIDEAN DIFFERENTIAL GEOMETRY

3. Euclidean duality

In this section the groundwork will be laid for the development of dual theories in euclidean differential geometry of surfaces. The necessary geometric characterizations will be made to secure a purely geometric basis for these theories.

Let x denote a vector from the origin to a point on the surface whose coordinates are $x^i(u^1, u^2)$, $(i = 1, 2, 3)$. The homogeneous coordinates x^1, x^2, x^3, 1 are specialized projective coordinates which can be obtained by selecting the three cartesian coordinate planes and the plane at infinity as the faces of the tetrahedron of reference, and by choosing the unit point so that the coordinates x^i, $(i = 1, 2, 3)$, are measured with respect to equal units on the three orthogonal axes. Let z^i, $(i = 1, 2, 3)$, denote the direction cosines of the normal to S at x and let z denote the point whose coordinates are z^i, $(i = 1, 2, 3, 4)$ wherein $z^4 = 0$. The pairs of functions (x, z) are solutions of a system of differential equations of the form (1.1), wherein

$$a_{ij} = c_i = e_i = 0, \qquad (i, j = 1, 2),$$

$$(3.1) \qquad (\Gamma^k_{ij}) = \{{}^k_{ij}\}$$

$$b^j_i = (g_{ir}d_{ik} - g_{kk}d_{ij})/g,$$

wherein $(i, j, k, r = 1, 2)$, with $r \neq i$, $k \neq j$, in which g_{ij} and d_{ij} are the coefficients of the first and second fundamental forms, respectively, $\{{}^k_{ij}\}$ are the Chris-

toffel symbols of the sedond kind for the first fundamental form of the surface, and g denotes the determinant of the quantities g_{ij}.

Two fundamental limits which will be called the *first* and *second ratios of a curve* C_λ *at a point* x will now be introduced. Let $(y - x)^2$ denote the square of the length of the vector from x to y. The quantity

$$\rho = \lim_{X \to x} (U^2 - x)^2/(U^1 - x)^2 \tag{3.2}$$

will be called the *first ratio* of C_λ at x. . Let $(y - x \mid z)$ denote the scalar product of the unit normal z and the vector from x to y. The quantity

$$\sigma = \lim_{X \to x} (U^2 - x \mid z)/(U^1 - x \mid z) \tag{3.3}$$

will be called *the second ratio of* C_λ *at* x.

The first ratio of C_λ at x is, clearly, given by

$$\rho = g_{22}\lambda^2/g_{11} \tag{3.4}$$

whenever the parametric net is not the minimal net.

If the minimal net is parametric, the differential equations of the surface take the form

$$\frac{\partial^2 x}{\partial u^{i2}} = \left(\frac{\partial(\log g_{12})}{\partial u^i}\right)\frac{\partial x}{\partial u^i} + d_{ii}z, \qquad \frac{\partial^2 x}{\partial u^1 \partial u^2} = d_{12}z, \tag{3.5}$$

$$\frac{\partial z}{\partial u^i} = -\left(d_{12}\frac{\partial x}{\partial u^i} + d_{ii}\frac{\partial x}{\partial u^j}\right)\Big/ g_{12}, \qquad (i, j = 1, 2), \quad i \neq j.$$

Making use of these equations together with the expansions (1.5) $(i = 1, 2)$, we obtain

$$\begin{aligned} U^i - x = {} & \frac{\partial x}{\partial u^i}\left(\frac{du^i}{dt}\Delta t + \cdots\right) + \frac{\partial x}{\partial u^j}\left(-d_{ii}\left(\frac{du^i}{dt}\right)^3 (\Delta t)^3/6g_{12} + \cdots\right) \\ & + z\left(d_{ii}\left(\frac{du^i}{dt}\right)^2 (\Delta t)^2/2 + \cdots\right), \quad (i, j = 1, 2), \quad i \neq j. \end{aligned} \tag{3.6}$$

Forming the scalar squares of the right member of (3.6) and simplifying, by making use of the relations

$$\left(\frac{\partial x}{\partial u^i}\right)^2 = 0, \quad (i = 1, 2), \qquad \left(\frac{\partial x}{\partial u^1} \,\middle|\, \frac{\partial x}{\partial u^2}\right) = g_{12}, \text{ we find}$$

$$(U^i - x)^2 = -(d_{ii})^2\left(\frac{du^i}{dt}\Delta t\right)^4 \Big/ 12 + \cdots, \qquad (i = 1, 2), \tag{3.7}$$

wherein $\dfrac{du^2}{dt} \Big/ \dfrac{du^1}{dt} = \lambda$. Consequently, if the minimal net is parametric

$$\rho = (d_{22})^2\lambda^4/(d_{11})^2. \tag{3.8}$$

If the parametric net is not the asymptotic net, it is clear that the second ratio of C_λ at x is given by

$$\sigma = d_{22}\lambda^2/d_{11}\,. \tag{3.9}$$

If the asymptotic net is parametric, the coefficients of the differential equations of S have the form (1.1) with $d_{11} = d_{22} = 0$ and $d_{12} \neq 0$. In the presence of these conditions the principal parts of the coefficients of z in the expansions (1.5) for $i = 1, 2$ are seen to be $\{^2_{1\,1}\}\, d_{12}\left(\frac{du^1}{dt}\,\Delta t\right)^3 \Big/ 6$ and $\{^1_{2\,2}\}\, d_{12}\left(\frac{du^2}{dt}\,\Delta t\right)^3 \Big/ 6$ respectively. Therefore, *if the asymptotic net is parametric, the second ratio of* C_λ *at* x *is given by*

$$\sigma = \{^1_{2\,2}\}\,\lambda^3/\{^2_{1\,1}\}. \tag{3.10}$$

The net $N_{\bar\lambda}$ derived from a family F_λ by the relation

$$\bar\rho(u^1, u^2) = 1/\rho(u^1, u^2), \tag{3.11}$$

wherein $\bar\rho$ and ρ are the first ratios of the curves $C_{\bar\lambda}$ and C_λ, respectively, consists of the integral curves of the differential equation

$$(du^2/du^1)^2 = (g_{11}/g_{22}\lambda)^2. \tag{3.12}$$

The directions at x *of the curves of* $N_{\bar\lambda}$ *which pass through* x, *clearly separate harmonically the directions at* x *of the parametric curves.* These directions are

$$du^2/du^1 = -g_{11}/g_{22}\lambda, \qquad du^2/du^1 = g_{11}/g_{22}\lambda. \tag{3.13}$$

It is clear that *if the parametric net is orthogonal, these are the directions of the tangents of* S *which bisect the angle formed by the tangents at* x *of the parametric curves.* The first equation of (3.13) can be obtained from equation (1.7) by replacing d_{11} and d_{22} by g_{11} and g_{22}, respectively. The tangent whose direction is defined by this equation will be called the R^*_λ*-correspondent of the tangent to* C_λ *at* x.

Analogously, the net N_{λ_1} derived from the family F_λ by the relation

$$\sigma_1(u^1, u^2) = 1/\sigma(u^1, u^2), \tag{3.14}$$

wherein σ_1 and σ are the second ratios of the curves C_{λ_1}, and C_λ, respectively, consists of the integral curves of the differential equation

$$(du^2/du^1)^2 = (d_{11}/d_{22}\lambda)^2. \tag{3.15}$$

The directions at x *of the curves of* N_{λ_1} *which pass through* x *separate harmonically the directions at* x *of the parametric curves.* These directions are given by

$$du^2/du^1 = -d_{11}/d_{22}\lambda, \qquad du^2/du^1 = d_{11}/d_{22}\lambda. \tag{3.16}$$

The first of these is the direction of the R_λ-correspondent of the tangent to C_λ at x.

Dual theories will be formulated in which the R_λ and R_λ^*-correspondents play fundamental roles. One of these theories may be obtained from the other by making use of the substitution

$$T = \begin{pmatrix} g_{ij} \\ d_{ij} \end{pmatrix}, \qquad (i, j = 1, 2). \tag{3.17}$$

In these theories the geometric entities of the following pairs are the duals of each other: the asymptotic net and the minimal net, an arbitrary conjugate net and a corresponding orthogonal net, the normal curvature for a direction at a point and the radius of normal curvature for the direction, the R_λ^*- and the R_λ-correspondents of the tangent to C_λ at x.

We present a simple euclidean geometric characterization for each of the R_λ^*- and R_λ-correspondents of a tangent to C_λ at x. These characterizations are suggested by the forms of the two equations

$$du^2/du^1 = -g_{11}/g_{22}\lambda \quad \text{and} \quad du^2/du^1 = -d_{11}/d_{22}\lambda.$$

Describe along the u^1-curve through x, the arcs $x\bar{U}_1^1$ and $x\bar{U}_2^1$ of lengths equal to that of the arc xU^2, and let $x\bar{U}_1^1$ denote the one of these arcs described in the sense of the arc xU^1. Describe along the u^2-curve through x, in the sense of the arc xU^2, the arc $x\bar{U}_1^2$ of length equal to that of the arc xU^1. The u^1-curve through $\bar{U}_1^2$ intersects the u^2-curves through $\bar{U}_1^1$ and $\bar{U}_2^1$ in the points which we denote by $\bar{X}_1$ and $\bar{X}$, respectively. As X tends to x along C_λ the points $\bar{X}$ and $\bar{X}_1$ tend to x along curves whose directions at x are given by the first and second equations of (3.13), respectively. *Thus the limit of the line joining x, $\bar{X}$ as X tends to x along C_λ is the R_λ^*-correspondent of the tangent to C_λ at x.*

Let U_1^1, U_2^1 denote the points on the u^1-curve through x at distances from the tangent plane to S at x numerically equal to that from the plane to the point U^2, and let the arc xU_1^1 be described in the sense of the arc xU^1. Describe along the u^2-curve through x, in the sense of the arc xU^2, the arc xU_1^2 whose endpoint U_1^2 is at a distance from the tangent plant to S at x equal to that from the plane to the point U^1. The u^2-curves through U_1^1 and U_2^1 intersect the u^1-curve through U_1^2 in the points which we denote by X_1 and X_2, respectively. As X tends to x along C_λ, the points X_1, X_2 tend to x along curves whose directions at x are given by (3.16). *The R_λ-correspondent of the tangent to C_λ at x is the limit of the line xX_2 or the limit of the line xX_1 according as the points U^1, U^2 are on the same side or opposite sides of the tangent plane to S at x.*

We are in a position now to state metric (euclidean) definitions and theorems which are the duals of the projective definitions and theorems of Chapter I. The dual of the principal associate net of a parametric net is the net whose directions at x are the double elements of the involution $\bar{\lambda} = -g_{11}/g_{22}\lambda$. This net will be called the *metric associate net* of the parametric net. The dual of Theorem (2.1) is

THEOREM (3.1) *At a general point x of S the tangent to an arbitrary curve C_λ*

through x and the R_λ^*-correspondent of this tangent are harmonic conjugates with respect to the tangents of the metric associate net of the parametric net.

The dual of Theorem (2.2) is

THEOREM (3.2) *The net of minimal curves is the metric associate net of an arbitrary orthogonal parametric net.*

The dual of the associate conjugate net of a parametric net will be called the *associate orthogonal net of the parametric net. This net is, clearly, the unique orthogonal net whose tangents at x separate harmonically the tangents of the parametric curves. Hence, the tangents at x of the associate orthogonal net of the parametric net bisect the angle formed by the tangents at x of the parametric curves.*

The dual of the R_λ-net will be called the R_λ^*-net. The tangents of this net at a point x are the tangent to C_λ at x and the R_λ^*-correspondent of this tangent. The duals of Theorems (2.4), (2.5), (2.6) and (2.7) may now be stated as follows:

THEOREM (3.3) *Every R_λ^*-net derived in association with an arbitrary orthogonal parametric net is an orthogonal net.*

THEOREM (3.4) *Corresponding to a parametric net which does not consist of minimal curves there is a unique R_λ^*-net whose tangents at a general point x separate harmonically the tangents to the parametric curves at the point. This R_λ^*-net is the associate orthogonal net of the parametric net.*

THEOREM (3.5) *The tangents at x to the parametric curves separate harmonically the R_λ^*-correspondent of the tangent at x to a minimal curve C_λ and the tangent to the other minimal curve at x.*

THEOREM (3.6) *If the minimal net is parametric, the limit of the line joining U^1, U^2 as X tends to x along C_λ is perpendicular to the tangent to C_λ at x.*

4. Euclidean characterizations of projectively defined nets and their duals by values of the first and second ratios

If a parametric net does not consist of asymptotic curves, the equations of its principal associate net and its associate conjugate net are $d_{11}(du^1)^2 + d_{22}(du^2)^2 = 0$ and $d_{11}(du^1)^2 - d_{22}(du^2)^2 = 0$, respectively. For such a parametric net the second ratio of a curve C_λ at x is given by $\sigma = d_{22}\lambda^2/d_{11}$. Hence we have immediately

THEOREM (4.1) *If a parametric net is not the asymptotic net, its principal associate net and its associate conjugate net are such that at a general point x the second ratios of the curves of these nets are equal to -1 and $+1$, respectively.*

Dually, we have

THEOREM (4.2) *If a parametric net is not the minimal net, its metric associate net and its associate orthogonal net are such that at a general point x the first ratios of the curves of these nets are equal to -1 and $+1$, respectively.*

5. The lines of curvature; euclidean characterizations of the curves of Segre and Darboux

On an unspecialized surface there exists an infinite class of nets each of which has the property that, relative to this net as parametric the equation

(5.1) $$\rho = \sigma$$

is satisfied independently of λ at a general point x of S. A net of this class will be called a duametric net. The necessary and sufficient condition that a parametric net which is not the asymptotic net nor the minimal net be a duametric net is found, by equating the right members of (3.4) and (3.9), to be

$$d_{11}g_{22} = d_{22}g_{11}\,. \tag{5.2}$$

Referred to an arbitrary parametric net, the lines of curvature are the integral curves of the equation

$$(g_{1i}d_{2j} - g_{2j}d_{1i})du^i du^j = 0, \qquad (i, j = 1, 2). \tag{5.3}$$

If the parametric net is a duametric net and the lines of curvature form a determinate net on S, equation (5.3) may be expressed in either of the forms

$$g_{11}(du^1)^2 - g_{22}(du^2)^2 = 0, \quad d_{11}(du^1)^2 - d_{22}(du^2)^2 = 0. \tag{5.4}$$

The forms of equations (5.3) and (5.4) are such that we have immediately

THEOREM (5.1) *The associate orthogonal net of an arbitrary duametric net is the same as its dual, the associate conjugate net. This self dual net is the net of lines of curvature. Along any line of curvature* $\rho = \sigma = 1$.

Since the definition of a duametric net is self dual, *the dual of a duametric net is a duametric net.* We shall reserve for special study in §(6) a particularly important pair of *dual duametric nets.*

If the minimal curves are parametric, the first ratio of C_λ at x is defined by (3.8). The curves characterized by putting $\rho = 1$ are the integral curves of the equation

$$(d_{22})^2(du^2)^4 - (d_{11})^2(du^1)^4 = 0. \tag{5.5}$$

Equation (5.3) reduces to $d_{22}(du^2)^2 = d_{11}(du^1)^2$ on putting $g_{11} = g_{22} = 0$, $g_{12} \neq 0$. Hence the curves (5.5) consist of the lines of curvature and the integral curves of the equation

$$d_{22}(du^2)^2 + d_{11}(du^1)^2 = 0. \tag{5.6}$$

The curves (5.6) may be shown to form the orthogonal net whose tangents at a general point x bisect the angles formed by the tangents to the lines of curvature at x. For reasons which will appear in §(6) the curves of this net will be called *the lines of mean curvature.* Thus we have

THEOREM (5.2) *If the minimal curves are parametric, there are two nets which are such that at a general point x of S the first ratios of the curves of these nets are equal to* 1. *These nets consist of the lines of mean curvature and the lines of curvature of S.*

Similarly we obtain

THEOREM (5.3) *If the asymptotic curves are parametric, a curve of S is a line of curvature if and only if at each of its points its first ratio is equal to* 1.

The dual of this theorem is

THEOREM (5.4) *If the minimal curves are parametric, a curve is a line of curvature if and only if at each of its points its second ratio is equal to* 1.

The curves of Segre and of Darboux are known to be the integral curves of the equations

$$(5.7) \qquad ({}_{2}^{1}{}_{2})(du^2)^3 - ({}_{1}^{2}{}_{1})(du^1)^3 = 0, \qquad ({}_{2}^{1}{}_{2})(du^2)^3 + ({}_{1}^{2}{}_{1})(du^1)^3 = 0$$

respectively, when the asymptotic curves are parametric. When the system of coordinates for x is cartesian $({}_{2}^{1}{}_{2}) = \{{}_{2}^{1}{}_{2}\}$ and $({}_{1}^{2}{}_{1}) = \{{}_{1}^{2}{}_{1}\}$. Hence we have immediately

THEOREM (5.5) *If the asymptotic curves are parametric, a curve of S is a curve of Segre or is a curve of Darboux according as its second ratio at each of its points is equal to* 1 *or* -1, *respectively.*

6. The orthogonal R_λ-net; the lines of mean curvature and the mean conjugate curves

If the parametric net is an arbitrarily selected non-asymptotic net, the differential equation of the R_λ-net is

$$(6.1) \qquad d_{11}\lambda(du^1)^2 + (d_{22}\lambda^2 - d_{11})du^1du^2 - d_{22}\lambda(du^2)^2 = 0.$$

The necessary and sufficient condition that this net be orthogonal is that the harmonic invariant of the two forms

$$g_{ij}du^idu^j, \qquad (i, j = 1, 2) \quad \text{and} \quad d_{11}\lambda(du^1)^2 + (d_{22}\lambda^2 - d_{11})du^1du^2 - d_{22}\lambda(du^2)^2$$

vanish, that is

$$(6.2) \qquad d_{11}g_{12} + (d_{11}g_{22} - d_{22}g_{11})\lambda - d_{22}g_{12}\lambda^2 = 0.$$

Thus there are in general two families F_λ for which the R_λ-net is orthogonal. However, we observe that both of these families belong to the same R_λ-net. Hence we have

THEOREM (6.1) *Corresponding to a non-asymptotic parametric net there is, in general, a unique orthogonal R_λ-net.*

Equations (6.2) may be regarded as the differential equation of this net.

However, there is a unique parametric net which, in marked contrast with a general parametric net, is such that every associated R_λ-net will be orthogonal. Equation (6.2) shows that this case occurs only if

$$(6.3) \qquad g_{12} = 0, \qquad d_{11}g_{22} - d_{22}g_{11} = 0.$$

This parametric net is clearly the *duametric net which is orthogonal.* Thus the directions at x of the curves of this net are those perpendicular directions which separate harmonically the principal directions at x. Hence these directions bisect the angles formed by the principal directions. The associated normal curvatures for the directions are, therefore, equal. These normal curvatures are, moreover, equal to the mean of the principal normal curvatures since the sum of the normal curvatures in any two perpendicular directions is equal to the sum of the principal normal curvatures at the point. In view of this property we shall call the curves of this orthogonal duametric net *the lines of mean curvature.* We state now

THEOREM (6.2) *The unique parametric net which is such that every associated R_λ-net is orthogonal is the orthogonal duametric net. The curves of this net are the lines of mean curvature.*

The equation for the net of lines of mean curvature when referred to an arbitrary parametric net, may be found by equating the right members of

$$1/r = d_{ij}du^i du^j/g_{ij}du^i du^j, \qquad (i, j = 1, 2),$$
$$(1/r_1 + 1/r_2)/2 = (g_{11}d_{22} + g_{22}d_{11} - 2g_{12}d_{12})/2g \tag{6.4}$$

wherein r is the radius of normal curvature of the direction du^2/du^1 and r_1, r_2 are the principal radii of normal curvature. The equation reduces to

$$(g_{11}B - g_{12}A)(du^1)^2 + (g_{11}C - g_{22}A)du^1du^2 + (g_{12}C - g_{22}B)(du^2)^2 = 0, \tag{6.5}$$

wherein A, B, C are the coefficients

$$A = 2(g_{11}d_{12} - g_{12}d_{11}), \qquad B = g_{11}d_{22} - g_{22}d_{11}, \qquad C = 2(g_{12}d_{22} - g_{22}d_{12}) \tag{6.6}$$

of equation (5.3) for the lines of curvature.

If $B = 0$ and $g_{12} \neq 0$, an arbitrary non-orthogonal duametric net is parametric. For such a choice of a parametric net equations (6.2) and (5.3) both reduce to $d_{11}(du^1)^2 = d_{22}(du^2)^2$. Hence we have

THEOREM (6.3) *The orthogonal R_λ-net derived in association with an arbitrary non-orthogonal duametric net consists of the lines of curvature of S.*

If the parametric net is not duametric, we find that the necessary and sufficient conditions that the orthogonal R_λ-net consist of the lines of curvature are

$$g_{11}d_{12} = g_{22}d_{12} = 0. \tag{6.7}$$

Hence we have

THEOREM (6.4) *If the parametric net is either the minimal net or a conjugate net and is not duametric, the associated orthogonal R_λ-net consists of the lines of curvature of S.*

The curves characterized by the properties which are the duals of those of the lines of mean curvature are found to be the mean conjugate curves. We recall that the mean conjugate curves are those for whose directions at a general point the radii of normal curvatures are equal to the mean of the principal radii of normal curvature at the point. The differential equation of the mean conjugate net is found to be

$$(d_{11}B - d_{12}A)(du^1)^2 + (d_{11}C - d_{22}A)du^1du^2 + (d_{12}C - d_{22}B)(du^2)^2 = 0, \tag{6.8}$$

wherein A, B, C are defined by (6.6). Since the net of lines of mean curvature is a duametric net, the mean conjugate net is likewise a duametric net. The mean conjugate net is parametric if and only if the equations

$$d_{12} = 0, \qquad d_{11}g_{22} - d_{22}g_{11} = 0, \tag{6.9}$$

are satisfied identically. The duals of Theorems (6.1) to (6.4) follow:

THEOREM (6.5) *Corresponding to a non-minimal parametric net there is in general, a unique conjugate R_λ^*-net.*

THEOREM (6.6) *The unique parametric net which is such that every associated R_λ^*-net is conjugate is the conjugate duametric net, known heretofore as the mean conjugate net.*

THEOREM (6.7) *The conjugate R_λ^*-net derived in association with an arbitrary non-conjugate duametric net consists of the lines of curvature of S.*

THEOREM (6.8) *If the parametric net is either the asymptotic net or an orthogonal net, the associated conjugate R_λ^*-net consists of the lines of curvature of S.*

7. The curvatures at x of the intersection of the surface S with its tangent plane at x

Let the asymptotic curves of S be parametric and let the points x, $\frac{\partial x}{\partial u^1}$, $\frac{\partial x}{\partial u^2}$ be the vertices of a *local triangle of reference at* x, with a unit point chosen so that any point whose general coordinates are given by an expression of the form

$$x_1 x + x_2 \frac{\partial x}{\partial u^1} + x_3 \frac{\partial x}{\partial u^2}$$

shall have local coordinates x_1, x_2, x_3. Let us introduce non-homogeneous coordinates by the definitions

$$\xi = x_2/x_1, \qquad \eta = x_3/x_1. \tag{7.1}$$

The intersection of S with its tangent plane at an ordinary point x is a curve C which has a node at x, the nodal tangents being the asymptotic tangents of S at x. In a neighborhood of x the equation for the branch of C which is tangent to the u^1-tangent is given by the power series development[4]

$$\eta = (_1{}^2{}_1)(\xi^2/3 - \varphi\xi^3/12) + \cdots, \tag{7.2}$$

wherein $\varphi = (_1{}^1{}_1) + \frac{\partial \log (_1{}^2{}_1)}{\partial u^1}$. Similarly, we have along the branch of C which is tangent to the u^2-tangent the development

$$\xi = (_2{}^1{}_2)(\eta^2/3 - \psi\eta^3/12) + \cdots, \tag{7.3}$$

wherein $\psi = (_2{}^2{}_2) + \frac{\partial \log (_2{}^1{}_2)}{\partial u^2}$.

Let the coordinates x be the homogeneous cartesian coordinates x_1, x_2, x_3, 1. Then $(_i{}^k{}_j)$, $(i, j, k = 1, 2)$, are defined by (3.1). Since $x_4 = 1$, the points $\frac{\partial x}{\partial u^1}$ and $\frac{\partial x}{\partial u^2}$ are the points at infinity on the u^1- and u^2-tangents of S at x. Let us alter the unit point so that new local coordinates $(\bar{\xi}, \bar{\eta})$ are measured with respect to equal units on the axes. Since $\frac{\partial x}{\partial u^1} \Big/ (g_{11})^{\frac{1}{2}}$ and $\frac{\partial x}{\partial u^2} \Big/ (g_{22})^{\frac{1}{2}}$ are unit vectors in

[4] Cf. P. D. G. page 73.

the directions of the u^1- and u^2-tangents to S at x, we must have the identity

$$x + \xi \frac{\partial x}{\partial u^1} + \eta \frac{\partial x}{\partial u^2} \equiv x + \bar{\xi} \frac{\partial x}{\partial u^1} \Big/ (g_{11})^{\frac{1}{2}} + \bar{\eta} \frac{\partial x}{\partial u^2} \Big/ (g_{22})^{\frac{1}{2}}. \tag{7.4}$$

Hence, we have the relations

$$\xi = \bar{\xi}/(g_{11})^{\frac{1}{2}}, \qquad \eta = \bar{\eta}/(g_{22})^{\frac{1}{2}}. \tag{7.5}$$

In terms of $\bar{\xi}$, $\bar{\eta}$ equations (7.2) and (7.3) take the respective forms

$$\bar{\eta} = \{{}^{2}_{1}{}_{1}\}\ (g_{22})^{\frac{1}{2}}\bar{\xi}^2/3g_{11} + \cdots, \tag{7.6}$$

$$\bar{\xi} = \{{}^{1}_{2}{}_{2}\}\ (g_{11})^{\frac{1}{2}}\bar{\eta}^2/3g_{22} + \cdots. \tag{7.7}$$

With respect to this oblique coordinate system defined by (7.5) the equation of a circle whose center is the point (h, k) and whose radius is r is found to be

$$(\bar{\xi} - h)^2 + (\bar{\eta} - k)^2 + 2(\bar{\xi} - h)(\bar{\eta} - k)g_{12}/(g_{11}g_{22})^{\frac{1}{2}} = r^2. \tag{7.8}$$

Demanding that this equation be satisfied by the series (7.6) identically in $\bar{\xi}$ as far as the terms of the second degree, we obtain the conditions that the circle (7.8) osculate at x the branch of C which is tangent to the u^1-tangent at x

$$h = -3(g_{11})^{\frac{1}{2}}g_{12}/2\{{}^{2}_{1}{}_{1}\}g, \quad k = 3(g_{11})^2(g_{22})^{\frac{1}{2}}/2\{{}^{2}_{1}{}_{1}\}g, \quad r^2 = 9(g_{11})^3/4\{{}^{2}_{1}{}_{1}\}^2g. \tag{7.9}$$

The substitution

$$T = \begin{pmatrix} g_{11} & g_{12} & \{{}^{2}_{1}{}_{1}\} & h & \bar{\xi} \\ g_{22} & g_{21} & \{{}^{1}_{2}{}_{2}\} & k & \bar{\eta} \end{pmatrix}$$

transforms equation (7.2) into equation (7.3) and equation (7.8) with h, k, r defined by (7.9) into an equation of form (7.8) with k, h, r defined by

$$k = -3(g_{22})^{\frac{1}{2}}g_{12}/2\{{}^{1}_{2}{}_{2}\}g, \qquad h = 3(g_{22})^2(g_{11})^{\frac{1}{2}}/2\{{}^{1}_{2}{}_{2}\}g, \tag{7.10}$$

$$r^2 = 9(g_{22})^3/4\{{}^{1}_{2}{}_{2}\}^2g.$$

Hence, *the circle* (7.8) *with* h, k, r *defined by* (7.10) *osculates at* x *the branch of* C *which is tangent to the* u^2*-tangent at* x.

The equation of the radical axis of the two circles which osculate C at x is readily found to be

$$\{{}^{2}_{1}{}_{1}\}(g_{22})^{\frac{1}{2}}\bar{\xi} - \{{}^{1}_{2}{}_{2}\}(g_{11})^{\frac{1}{2}}\bar{\eta} = 0. \tag{7.11}$$

In terms of the original local coordinates ξ, η this equation is

$$\{{}^{2}_{1}{}_{1}\}g_{22}\xi - \{{}^{1}_{2}{}_{2}\}g_{11}\eta = 0. \tag{7.12}$$

The direction of this line is given by

$$\eta/\xi = g_{22}\{{}^{2}_{1}{}_{1}\}/g_{11}\{{}^{1}_{2}{}_{2}\}.$$

The direction of the R_λ-correspondent of the tangent at x in either principal direction is given by

$$du^2/du^1 = -g_{22}\{{}^{2}_{1\,1}\}/g_{11}\{{}^{1}_{2\,2}\}$$

Hence we have

THEOREM (7.1) *If the asymptotic curves are the parametric curves of S, the R_λ-correspondent of the tangent at x in either principal direction and the radical axis of the osculating circles of C at x are conjugate tangents.*

On observing the forms (3.4) and (3.10) for the first and second ratios of a curve C_λ at x we find that the only non-asymptotic direction λ for which these ratios are equal is that defined by

$$\lambda = g_{22}\{{}^{2}_{1\,1}\}/g_{11}\{{}^{1}_{2\,2}\}. \tag{7.13}$$

Hence we have

THEOREM (7.2) *If a curve C_λ is not tangent to an asymptotic curve at x, its tangent line at x is the radical axis of the osculating circles of C at x if and only if first and second ratios of C_λ at x are equal.*

Let K_1K_2 denote the curvatures of the circles which osculate at x the branches of C which are tangent to the u^1- and u^2-tangents respectively. These curvatures will be called the *curvatures of tangential section of S at x*. We have in view of (7.9) and (7.10)

$$K_1 = 2e_1\{{}^{2}_{1\,1}\}g^{\frac{1}{2}}/3(g_{11})^{\frac{3}{2}}, \qquad k_2 = 2e_2\{{}^{1}_{2\,2}\}g^{\frac{1}{2}}/3(g_{22})^{\frac{3}{2}}, \tag{7.14}$$

wherein $g^{\frac{1}{2}} = (g_{11}g_{22} - g_{12}^2)^{\frac{1}{2}}$ and $e_1 = \pm 1$, $e_2 = \pm 1$, the signs of e_1 and e_2 being chosen so that K_1 and K_2 are positive. These curvatures are equal if and only if

$$(g_{11}/g_{22})^{\frac{1}{2}} = e(\{{}^{2}_{1\,1}\}/\{{}^{1}_{2\,2}\})^{\frac{1}{2}}, \tag{7.15}$$

with $e = \pm 1$ the sign being selected so that the right member is positive. The left member of (7.15) is a principal direction and the right member is a direction of Segre or a direction of Darboux according as e is $+1$ or -1 respectively. Substituting in (7.13) the expression obtained for g_{22}/g_{11} from (7.15), we find that (7.13) becomes

$$\lambda = (\{{}^{2}_{1\,1}\}/\{{}^{1}_{2\,2}\})^{\frac{1}{2}}. \tag{7.16}$$

Thus we find that under conditions (7.15) the direction defined by (7.13) is a direction of Segre and also, in consequence of (7.15), a principal direction. Moreover, if the direction defined by (7.13) is a direction of Segre, equations (7.15) follow. We may therefore state

THEOREM (7.3) *The necessary and sufficient condition that the curvatures of tangential section of S at x be equal is that the direction of the radical axis of the osculating circles of C at x be a direction of Segre. The condition is, moreover,*

necessary and sufficient that this direction of Segre and its conjugate direction of Darboux be principal directions at x.

In view of Theorems (5.3), (5.5) and (7.3) we state in conclusion

COROLLARY (7.4) *If the asymptotic curves are parametric, the one parameter family F_λ of curves enveloped by the radical axis of the osculating circles of C at x as x varies over S has the property that the first and second ratios of C_λ at x are both equal to 1 if and only if the curvatures of tangential section of S at x are equal.*

PRINCETON UNIVERSITY
UNIVERSITY OF KANSAS

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

ON FAMILIES OF MUTUALLY EXCLUSIVE SETS

BY P. ERDÖS AND A. TARSKI

(Received August 11, 1942)

In this paper we shall be concerned with a certain particular problem from the general theory of sets, namely with the problem of the existence of families of mutually exclusive sets with a maximal power. It will turn out—in a rather unexpected way—that the solution of these problems essentially involves the notion of the so-called "inaccessible numbers." In this connection we shall make some general remarks regarding inaccessible numbers in the last section of our paper.

§1. FORMULATION OF THE PROBLEM. TERMINOLOGY[1]

The problem in which we are interested can be stated as follows: Is it true that every field $\mathfrak{F}$ of sets contains a family of mutually exclusive sets with a maximal power, i.e. a family $\mathfrak{G}$ whose cardinal number is not smaller than the cardinal number of any other family $\mathfrak{H}$ of mutually exclusive sets contained in $\mathfrak{F}$.

By a field of sets we understand here as usual a family $\mathfrak{F}$ of sets which together with every two sets X and Y contains also their union $X \cup Y$ and their difference $X - Y$ (i.e. the set of those elements of X which do not belong to Y) among its elements. A family $\mathfrak{G}$ is called a family of mutually exclusive sets if no set X of $\mathfrak{G}$ is empty and if any two different sets of $\mathfrak{G}$ have an empty intersection.

A similar problem can be formulated for other families e.g. for rings of sets, i.e. for families which together with any two sets X and Y also contain their union $X \cup Y$ and their intersection $X \cap Y$ among their elements. We obtain an especially interesting particular case of this problem by referring it to the ring of open sets of a topological space S with power $2^{\aleph_0}$.

It turns out that the solution of our problem is in general positive; however it is negative in certain exceptional cases. To examine the problem thoroughly we must first subject it to a certain transformation by using some notions from the arithmetic of cardinal numbers.

We shall denote the cardinal number (or power) of a set S by $\mathfrak{c}(S)$.

A cardinal number $\mathfrak{n}$ is called a limit number if $\mathfrak{n} \neq 0$ and if among the cardinal numbers $\mathfrak{x} < \mathfrak{n}$ there is no largest one. The number $\mathfrak{n}$ is called *singular* if it can be expressed as a sum of less than $\mathfrak{n}$ numbers $\mathfrak{m}$, each of which is smaller than $\mathfrak{n}$.

[1] For the concepts and results of the general theory of sets, which are applied in this paper, see Hausdorff, Mengenlehre: however as regards the concept of an inaccessible number cf. Tarski, Über unerreichbare Kardinalzahlen, Fund. Math. Vol. 30 (1938) p. 68–89. For the concepts and results from the theory of partially ordered sets, lattices, Boolean algebras, etc., see G. Birkhoff, Lattice theory. For topological concepts see Kuratowski, Topologie I.

If such a representation is impossible the number $\mathfrak{n}$ is called *regular*. Regular limit numbers are also referred to as "inaccessible" or "weakly inaccessible" numbers.

As is well known, every limit number is an infinite number, and every singular infinite number is a limit number. The problem of the existence of regular limit numbers $> \aleph_0$ is thus far unsolved, and presumably will never be solved on the basis of the axiom systems upon which the general theory of sets is constructed at present. At any rate the existence of the numbers in question cannot be derived from these axiom systems provided they are consistent; on the other hand it seems highly improbable that these systems cease to be consistent if we enrich them by adding new existential axioms which secure the existence of the inaccessible numbers.

$\mathfrak{F}$ being a family of sets let us denote by $\mathfrak{d}(\mathfrak{F})$ the smallest cardinal number which is $> \mathfrak{c}(\mathfrak{G})$ for every family $\mathfrak{G}$ of mutually exclusive sets contained in F. If $\mathfrak{d}(\mathfrak{F})$ is not a limit number, the family F obviously contains a subfamily $\mathfrak{G}$ of mutually exclusive sets with a maximal power. Thus our problem reduces now to the following one:

$\mathfrak{n}$ being a limit number is it true that for every field (or ring) of sets we have $\mathfrak{d}(\mathfrak{F}) \neq \mathfrak{n}$?

We shall show that the solution of this problem depends on the properties of the number $\mathfrak{n}$: the answer is affimative if $\mathfrak{n}$ either $= \aleph_0$ or is a regular number (Theorem 1), is negative only for the hypothetical regular limit numbers $> \aleph_0$ (Theorem 2). If in particular the problem is applied to the ring of all open sets of a topological space[2] with $\mathfrak{c}(S) = 2^{\aleph_0}$, then its positive solution proves to be equivalent with the statement that there is no inaccessible number $> \aleph_0$ and $\leqq 2^{\aleph_0}$ (Corollary 3).

In order to formulate the positive part of our result in as general form as possible, we shall use the terminology of partially ordered sets.

Let S be an arbitrary set which is partially ordered by the binary relation $\leqq$. If x is an element of S, we write $S(x)$ to denote the partially ordered set of all elements $y \in S$ which are $\leqq x$. The symbol Λ will denote a null element of S, i.e. an element x such that $x \leqq y$ for every $y \in S$. Two elements y and z of S are called disjoint if $y \neq \Lambda$, $z \neq \Lambda$ and if for every $x \in S$ the formulas $x \leqq y$ and $x \leqq z$ imply $x = \Lambda$. We do not here assume that the partially ordered set necessarily contains a null element. In fact without loss of generality we could confine ourselves to the consideration of sets which do not contain such elements; and in this case we could simply say that two elements y and z are called disjoint if there is no element x such that $x \leqq y$ and $x \leqq z$.

A subset T of a partially ordered set S such that every element of T is $\neq \Lambda$ and every two different elements of T are disjoint is called a set of mutually

[2] By topological spaces we mean here the spaces with the closure operation satisfying the axioms I–III of Kuratowski (op. cit. p. 77). However the space which will be constructed in the proof of Corollary 2 will also satisfy axiom IV (normality) (pp. 95–101, ibid.).

exclusive elements. Again we denote by $\mathfrak{d}(\mathfrak{S})$ the smallest cardinal number $> \mathfrak{c}(T)$ for every $T \subseteq S$ of sets of mutually exclusive elements; moreover we write for every element $x \in T$

$$\mathfrak{d}(x) = \mathfrak{d}(\mathfrak{S}(x))$$

In view of this formula $\mathfrak{d}$ constitutes an example of a function f which correlates with every element of a partially ordered set a cardinal number $f(x)$. This function is obviously increasing, for we have

$$\mathfrak{d}(x) \leqq \mathfrak{d}(y)$$

for every two elements x and y such that $x \leqq y$. Many other examples of this kind of increasing functions are also known; e.g. the function c defined for every $x \in S$ by the formula

$$\mathfrak{c}(x) = \mathfrak{c}(S(x)).$$

Still another example is constituted by the function g defined in the following way: for every $x \in S$, $g(x)$ is the smallest cardinal number $\mathfrak{n}$ such that there is a basis B of the set $S(x)$ with power $c(B) = \mathfrak{n}$; by a basis we here understand a set $B \subseteq S(x)$ such that every element of $S(x)$ is the union (the least upper bound) of elements of B. To every increasing function f of the kind considered there corresponds a certain notion of *homogeneity* of partially ordered sets. We say generally that an element x of a partially ordered set S is homogeneous with respect to an increasing function f, which is defined over the set S and assumes cardinal numbers as values, or simply that x is f-homogeneous, if $x \neq \Lambda$ and if $f(x) = f(y)$ for every element $y \in S$ such that $y \neq \Lambda$ and $y \leqq x$. If the set S contains a unit element u i.e. an element x such that $y \leqq x$ for every $y \in S$), and if u is f-homogeneous, the whole set S is called f-homogeneous.

SOLUTION OF THE PROBLEM

We shall begin with two simple lemmas concerning f-homogeneous elements

LEMMA 1. *Let S be a partially ordered set, and f an increasing function which correlates with every element $x \in S$ a cardinal number $f(x)$. Then for every element $x \neq \Lambda$ there exists an f-homogeneous element $y \leqq x$.*

PROOF. Consider all the cardinal numbers $f(y)$ correlated with the elements $y \leqq x$, $y \neq \Lambda$. Among these cardinal numbers there certainly exists a smallest, say $\mathfrak{n}$ (by the well ordering theorem); and it is easily seen that every element such that

$$y \leqq x, \qquad f(y) = \mathfrak{n}$$

is f-homogeneous.

LEMMA 2. *Under the hypothesis of* LEMMA 1. *there exists a set $T \subseteq S$ of mutually exclusive f-homogeneous elements such that no element of S is disjoint with all elements of T.*

PROOF. It can be easily shown (e.g. with the help of well ordering) that there exists a maximal set T of mutually exclusive f-homogeneous elements of S; i.e., a set T of mutually exclusive f-homogeneous elements of S which is not a proper subset of any other set with the same property. Hence by LEMMA 1 it follows that no element of S—whether homogeneous or not—is disjoint with every element of T, q.e.d.

As an immediate consequence of LEMMA 2 we obtain the following theorem which, however, will not be applied in this paper.

Let B be a Boolean algebra, and f an increasing function which correlates with every element x of B a cardinal number $f(x)$. Then every element of B—and, in particular, the unit element—can be represented as the union of mutually exclusive f-homogeneous elements of B; and therefore B is isomorphic with a direct sum of f-homogeneous Boolean algebras.

The following three lemmas will lead us directly to THEOREM 1, which is one of the main results of this paper.

LEMMA 3. *If S is a partially ordered set and $\mathfrak{d}(S)$ is a limit number, then S contains a $\mathfrak{d}$-homogeneous element x with $\mathfrak{d}(x) = \mathfrak{d}(S)$.*

PROOF. Assume that, on the contrary, S does not contain a $\mathfrak{d}$-homogeneous element x with $\mathfrak{d}(x) = \mathfrak{d}(S)$. By applying LEMMA 2 to the function $f = \mathfrak{d}$ we obtain a set $T \subseteq S$ of mutually exclusive $\mathfrak{d}$-homogeneous elements, with the property that no element of S is disjoint with every element of T. According to our assumption we have:

$$\mathfrak{d}(t) < \mathfrak{d}(S) \quad \text{for every element} \quad t \in T; \tag{1}$$

moreover, the definition of $\mathfrak{d}$ implies:

$$\mathfrak{c}(T) < \mathfrak{d}(S) \tag{2}$$

Since $\mathfrak{d}(S)$ is an infinite cardinal number, we have

$$(\mathfrak{d}(S))^2 = \mathfrak{d}(S); \tag{3}$$

hence, by (1) and (2), we obtain:

$$\sum_{t \in T} \mathfrak{d}(t) \leqq \mathfrak{c}(T) \cdot \mathfrak{d}(S) \leqq (\mathfrak{d}(S))^2 = \mathfrak{d}(S). \tag{4}$$

We want now to show that in the latter formula '$\leqq$' may be replaced by '='. In fact, consider an arbitrary set $U \subseteq S$ of mutually exclusive elements. As was mentioned before, no element of U can be disjoint with every element of T. Hence (by using the axiom of choice) we can correlate with every element $u \in U$ first an element $t_u \in T$, and then an element $v_u \in S$ such that $v_u \neq \Lambda$, $v_u \leqq u$, and $v_u \leqq t_u$. Let V be the set of all these elements v_u. It can easily be seen that the correspondence between the elements of U and those of V is one-to-one, and that therefore the sets U and V have the same power. If, on the other hand, we denote by V_t (where t is a given element of T) the set of all those elements v_u which are $\leqq t$, we see at once that V is the union of all these sets V_t;

$$V = \bigcup V_t$$

and that

$$\mathfrak{c}(V_t) < \mathfrak{d}(t) \quad \text{for every} \quad t \in T.$$

Hence

$$\mathfrak{c}(U) = \mathfrak{c}(V) \leqq \sum_{t \in T} \mathfrak{d}(t).$$

Since the latter formula holds for every set $U \subseteq S$ of mutually exclusive elements, we infer from the definition of $\mathfrak{d}$ that either

$$\mathfrak{d}(S) \leqq \sum_{t \in T} \mathfrak{d}(t) \tag{5}$$

or else $\mathfrak{d}(S)$ is the cardinal number which immediately follows $\sum_{t \in T} \mathfrak{d}(t)$. However, the second alternative is excluded, $\mathfrak{d}(S)$ being by hypothesis a limit number, and therefore (5) holds. The formulas (4) and (5) give at once:

$$\sum_{t \in T} \mathfrak{d}(t) = \mathfrak{d}(S). \tag{6}$$

From (1), (2), (3), and (6) it follows that for every cardinal number $\mathfrak{x} < \mathfrak{d}(S)$ there exists an element $t \in T$ such that $\mathfrak{x} < \mathfrak{d}(t)$. For if we had:

$$\mathfrak{d}(S) > \mathfrak{x} \geqq \mathfrak{d}(t) \quad \textit{for every} \quad t \in T$$

we should have, by (1), (2), and (3),

$$\sum_{t \in T} \mathfrak{d}(t) \leqq c(T)\mathfrak{x} < (\mathfrak{d}(S))^2 = \mathfrak{d}(S)$$

which obviously contradicts (6). Hence we can easily construct (with the help of the well ordering theorem) a well ordered transfinite sequence of elements $t_0, t_1, \cdots, t_\xi, \cdots \in T$ of an ordinal type τ which satisfy the following conditions:

(7) $\mathfrak{d}(t_{\xi_1}) < \mathfrak{d}(t_{\xi_2})$ for $\xi_1 < \xi_2 < \tau$, τ being a limit ordinal number,

and

$$\sum_{\xi < \tau} \mathfrak{d}(t_\xi) = \mathfrak{d}(S). \tag{8}$$

Consider an arbitrary ordinal number $\xi < \tau$. By (7) we have

$$\mathfrak{d}(t_\xi) < \mathfrak{d}(t_{\xi+1})$$

Hence by virtue of the definition of $\mathfrak{d}$, there exists a set $W_\xi \subseteq S(t_{\xi+1})$ of mutually exclusive elements with power:

$$\mathfrak{c}(W_\xi) = \mathfrak{d}(t_\xi) \quad \text{for every} \quad \xi < \tau. \tag{9}$$

Putting

$$W = \bigcup_{\xi < \tau} W_\xi \tag{10}$$

we easily see that W is a set of mutually exclusive elements of S, the sets $W_0, W_1, \cdots, W_\xi, \cdots$ being mutually exclusive. We have moreover, by (8), (9), and (10),

$$\mathfrak{c}(W) = \mathfrak{d}(S).$$

In this way we have arrived at a contradiction; for $\mathfrak{d}(S)$ is by definition $> \mathfrak{c}(X)$ for every set $X \subseteq S$ of mutually exclusive elements. Thus we must reject our original supposition, and assume that S has a d-homogeneous element x with $\mathfrak{d}(x) = \mathfrak{d}(S)$, q.e.d.

LEMMA 4. *If x is a $\mathfrak{d}$-homogeneous element of a partially ordered set S, then* $\mathfrak{d}(x) \neq \aleph_0$.

PROOF. Assume $\mathfrak{d}(x) = \aleph_0$. By the definition of $\mathfrak{d}$ there exist two disjoint elements x_1 and x_2 which are $\leqq x$. Since x is $\mathfrak{d}$-homogeneous, we have further $\mathfrak{d}(x_2) = \aleph_0$, and therefore there exist two disjoint elements, $x_{2,1}$ and $x_{2,2}$ which are $\leqq x_2 \leqq x$. By continuing this procedure indefinitely we obtain (with the help of the axiom of choice) an infinite sequence of mutually exclusive elements $x_1, x_{2,1}, x_{2,2,1}, \cdots$ which are all $\leqq x$; but this clearly contradicts our assumption. Hence $\mathfrak{d}(x) \neq \aleph_0$, q.e.d.

LEMMA 5. *If x is a $\mathfrak{d}$-homogeneous element of a partially ordered set, then* $\mathfrak{d}(x)$ *is not a singular limit number.*

PROOF. Assume, on the contrary, that $\mathfrak{d}(x)$ is a singular limit number. Thus $\mathfrak{d}(x)$ can be represented in the form:

$$\mathfrak{d}(x) = \sum_{i \in C} \mathfrak{m}_i \tag{11}$$

where C is a certain set of power $< \mathfrak{d}(x)$, and every number $\mathfrak{m}_i$ is also $< \mathfrak{d}(x)$. Since $\mathfrak{c}(C) < \mathfrak{d}(x)$, we can correlate with every element $i \in C$ an element $x_i \leqq x$ in such a way that any two elements x_{i_1} and x_{i_2} $(i_1 \neq i_2)$ are disjoint. The element x being $\mathfrak{d}$-homogeneous, we have for every element x_i:

$$\mathfrak{d}(x_i) = \mathfrak{d}(x) > \mathfrak{m}_i;$$

consequently we can correlate with every element x_i a set $T_i \subseteq S(x_i)$ of mutually exclusive elements with power $\mathfrak{c}(T_i) = \mathfrak{m}_i$. Hence in view of (11) it is easily seen that the set T defined by means of the formula

$$T = \bigcup_{i \in C} T_i$$

is a set of mutually exclusive elements of $S(x)$ with power

$$\mathfrak{c}(T) = \sum_{i \in C} \mathfrak{M}_i = \mathfrak{d}(x).$$

But this is impossible, since $\mathfrak{d}(x)$ must be by definition $> \mathfrak{c}(T)$. Thus we must assume that $\mathfrak{d}(x)$ is not a singular limit number, q.e.d.

Lemmas 3, 4, and 5 imply directly

THEOREM 1. *If $\mathfrak{n}$ is either equal to $\aleph_0$ or is a singular limit number then there is no partially ordered set S such that* $\mathfrak{d}(S) = \mathfrak{n}$.

REMARK. Theorem 1 applies directly to various special partially ordered sets, e.g., to lattices or Boolean algebras. It can also be applied to an arbitrary family $\mathfrak{F}$ of sets, for every such family is partially ordered by the relation of inclusion $\subseteq$. It should be noticed, however, that two sets of a family $\mathfrak{F}$ which are disjoint from the point of view of the theory of partially ordered sets are not necessarily disjoint in the usual set-theoretic meaning. On the other hand, it is easily seen that the two meanings of the notion of disjointness coincide if the family $\mathfrak{F}$ contains the empty set among its elements and if, with any two sets X and Y belonging to $\mathfrak{F}$ their intersection $X \cap Y$ also belongs to $\mathfrak{F}$. Thus Theorem 1 applies literally to every field of sets, and even to every ring of sets which contains the empty set, e.g., to the ring of all open sets of a topological space; and it can be very easily shown that the theorem holds for an arbitrary ring of sets, even if it does not contain the empty set (for in this case the ring does not contain any two sets which are disjoint in the set-theoretic sense).

THEOREM 2. *If $\mathfrak{n}$ is a regular cardinal number $> \aleph_0$, then for every set S of power $\mathfrak{n}$ there exists a field $\mathfrak{F}$ of subsets of S such that $\mathfrak{c}(F) = \mathfrak{d}(\mathfrak{F}) = n$.*

PROOF. We could assume that $\mathfrak{n}$ is a limit number, for otherwise the proof presents no difficulty; however, no use will be made here of this assumption.

We shall first prove the theorem for a particular set N of power $\mathfrak{n}$, which will be defined as follows. Let us write for every ordinal number α:

(1) $\mathfrak{c}(\alpha)$ = the power of the set of all ordinal numbers $\xi < \alpha$. By the well-ordering theorem there exists an ordinal number ν such that

(2) $\mathfrak{c}(\nu) = \mathfrak{n}$, while $\mathfrak{c}(\xi) < \mathfrak{n}$ for every number $\xi < \nu$.

(3). N = the set of all transfinite sequences σ of ordinal numbers $\sigma_0, \sigma_1, \cdots$, which satisfy the following conditions:

(i) $\sigma_\xi \leqq \xi$ for every number $\xi < \nu$;

(ii) there are only finitely many numbers $\xi < \nu$ such that $\sigma_\xi \neq 0$.

Since

(4). $n = n^0 + n^1 + \cdots + n^k + \cdots \qquad (k < \aleph_0)$,

it is easily seen from (1), (2), and (3) that N has in fact power $\mathfrak{n}$.

We are now going to correlate to every number $\xi < \nu$ a family $\mathfrak{H}$ of subsets X of N so as to satisfy the following conditions:

(5). $\mathfrak{H}_\xi$ is a family of mutually exclusive sets, and $\bigcup_{X \in \mathfrak{H}_\xi} X = N$

(6). $\mathfrak{c}(\mathfrak{H}_\xi) = \mathfrak{c}(\xi + 1) < \mathfrak{n}$;

(7). if $\xi_1, \xi_2, \cdots \xi_n$ is any finite sequence of distinct ordinal numbers $< \nu$, and $X_1, X_2, \cdots, X_n$ any finite sequence of sets such that $X_1 \in \mathfrak{H}_{\xi_1}$, $X_2 \in \mathfrak{H}_{\xi_2}, \cdots, X_n \in \mathfrak{H}_{\xi_n}$, then the intersection $X_1 \cap \cdots \cap X_n$ is not empty.

To obtain such families $\mathfrak{H}_\xi$ we put:

(8). $N_{\xi,\eta}$ = the set of all sequences $\sigma \in N$ such that $\sigma_\xi = \eta$, $\eta \leqq \xi < \nu$

(9). $\mathfrak{H}_\xi$ = the family of all sets $N_{\xi,0}, N_{\xi,1} \cdots N_{\xi,\eta} \cdots$ where $\eta \leqq \xi < \nu$.

The proof that the families $\mathfrak{H}_\xi$ thus defined satisfy the conditions (5), (6), and (7) does not present any difficulties.

Finally we construct the field $\mathfrak{F}$ by putting

(10). $\mathfrak{H} = \bigcup_{\xi < \nu} \mathfrak{H}_\xi$

(11). $\mathfrak{F}$ = the smallest field of sets which contains all the sets of $\mathfrak{H}$; or, in other words,: $\mathfrak{F}$ = the family of all sets which are finite unions of finite intersections of sets $X \in \mathfrak{H}$ and their complements $N - X$.

We shall prove that $\mathfrak{F}$ satisfies the conclusion of our theorem. If $\mathfrak{n}$ is an infinite number, it is easily seen from (2), (6), and (10) that the family $\mathfrak{H}$ has power $\mathfrak{n}$. Hence by (4) and (11) it follows that $\mathfrak{F}$ has also power $\mathfrak{n}$. Furthermore, (2), (5), (6), (10), and (11) imply that, for every number $\mathfrak{x} < \mathfrak{n}$, $\mathfrak{F}$ does contain $\mathfrak{x}$ mutually exclusive sets. Hence we have

(12). $\mathfrak{c}(\mathfrak{F}) = \mathfrak{n} \leqq \mathfrak{d}(\mathfrak{F})$.

It remains to show that every family $\mathfrak{G} \subseteq \mathfrak{F}$ of mutually exclusive sets has a power $< n$. We shall show it first for families of a rather special character.

Let us agree to say that a set $X \in \mathfrak{F}$ is of the l^{th} order (where l is any positive integer) if it is not empty and can be represented as an intersection of l different sets of the family $\mathfrak{H}$. We are going to establish certain simple properties of sets of the l^{th} order.

(13). Every set X of the l^{th} order can be represented uniquely in the form

$$X = X_1 \cap \cdots \cap X_l$$

where $X_1 \in \mathfrak{H}_{\xi_1}, \cdots, X_l \in \mathfrak{H}_{\xi_l}$, and $\xi_i < \xi_k < \nu$ for $1 \leqq i < k \leqq l$.

In fact, the possibility of such a representation follows directly from the definition of the sets of the l^{th} order; two different sets X_i and X_k cannot belong to the same family $\mathfrak{H}$, for by (5) the set $X \subseteq X_i \cap X_k$ would be then empty. Assume that the set X has two representations of this kind:

$$X = X_1 \cap \cdots \cap X_l = Y_1 \cap \cdots \cap Y_l$$

where $X_i \in \mathfrak{H}_{\xi_i}$, $Y_i \in H_{\eta_i}$, $\xi_i < \xi_k < \nu$, and $\eta_i < \eta_k < \nu$ for $1 \leqq i < k \leqq l$. If these representations are different, at least one of the sets $X_1, \cdots, X_l$, let us say X_i, cannot occur among the sets $Y_1, \cdots, Y_l$; and similarly a certain set Y_j cannot occur among the sets $X_1, \cdots, X_l$. Hence the number ξ_i must be different from each of the numbers $\eta_1, \cdots, \eta_l$. For, if we had $\xi_i = \eta_k$ $(1 \leqq k \leqq l)$, the sets X_i and Y_k $(X_i \neq Y_k)$ would belong to the same class $\mathfrak{H}_{\xi_i} = \mathfrak{H}_{\eta_k}$; and therefore by (5) the set $X \subseteq X_i \cap Y_k$ would be empty. For the same reason the number η_j must be different from each of the numbers $\xi_1, \cdots, \xi_l$. Thus, in particular, $\xi_i \neq \eta_j$, and therefore at least one of these two numbers is $\neq 0$. Assume, e.g., $\xi_i \neq 0$. By (6) there is a set $X_i' \in \mathfrak{H}_{\xi_i}$ which is different from X_i. By (5) the sets X_i and X_i' are disjoint; consequently the intersection

$$X_i' \cap X = X_i' \cap X_1 \cap \cdots \cap X_l$$

is empty. Hence the intersection

$$X_i' \cap X = X_i' \cap Y_1 \cap \cdots \cap Y_l$$

must also be empty; but this clearly contradicts (7), all the numbers $\xi_i, \eta_1, \eta_2, \cdots \eta_l$ being distinct. Thus the two representations of X cannot be different.

In what follows we shall refer to the sets $X_1, \cdots, X_l$ occurring in the representation (13) of a set X as the *factors* of X.

(14). In order that two sets X and Y of the l^{th} order be disjoint it is necessary and sufficient that they have two factors X_i and Y_i which are different, but belong to the same family $\mathfrak{H}$.

This follows directly from (5), (7), and (13).

(15). If two disjoint sets X and Y of the $(l+1)^{\text{th}}$ order have a common factor $X_i = Y_j$, and if X^* and Y^* are the intersections of the remaining factors of X and Y respectively, then X^* and Y^* are disjoint sets of the l^{th} order.

This can be easily obtained from (13) and (14).

Now we can prove by induction with respect to l:

(16). Every family $\mathfrak{G} \subseteq \mathfrak{F}$ of mutually exclusive sets of the l^{th} order has power $< \mathfrak{n}$.

(16) is clearly true for $l = 1$. In fact, in this case $\mathfrak{G}$ is contained in the family

$$\mathfrak{H} = \bigcup_{\xi < \nu} \mathfrak{H}_\xi .$$

If there were two sets X_1 and X_2 of $\mathfrak{G}$ which belonged to two different families $\mathfrak{H}_{\xi_1}$ and $\mathfrak{H}_{\xi_2}$, they would not be disjoint, on account of (7). Therefore there must be a $\xi < \nu$ such that $\mathfrak{G} \subseteq \mathfrak{H}_\xi$, and hence, by (6), $\mathfrak{c}(\mathfrak{G}) < \mathfrak{n}$.

Now assume that (16) has been proved for a given positive integer l, and consider a family $\mathfrak{G} \subseteq \mathfrak{F}$ of mutually exclusive sets of the $(l+1)$st order. Let X be any set of $\mathfrak{G}$, and let $\mathfrak{H}_{\xi_1}, \mathfrak{H}_{\xi_2}, \cdots, \mathfrak{H}_{\xi_{l+1}}$ $(\xi_1 < \xi_2 < \cdots < \xi_{l+1} < \nu)$ be those families $\mathfrak{H}_\xi$ which by (13) contain among their elements a factor of X. By (14) every set of $\mathfrak{G}$ must have a factor belonging to at least one of the families $\mathfrak{H}_{\xi_1}, \cdots, \mathfrak{H}_{\xi_{l+1}}$, and thus also to their union

$$\mathfrak{H}^* = \mathfrak{H}_{\xi_1} \cup \cdots \cup \mathfrak{H}_{\xi_{l+1}} . \tag{17}$$

For every set Z of H^* let us denote by $\mathfrak{G}(Z)$ the family of all those sets $Y \in \mathfrak{G}$ which have Z as a factor. We have thus a decomposition of $\mathfrak{G}$ in subfamilies $\mathfrak{G}(Z)$ (which are not unnecessarily mutually exclusive):

$$\mathfrak{G} = \bigcup_{Z \in \mathfrak{H}^*} \mathfrak{G}(Z).$$

Hence

$$\mathfrak{c}(\mathfrak{G}) \leqq \sum_{Z \in \mathfrak{H}^*} \mathfrak{c}[\mathfrak{G}(Z)]. \tag{18}$$

Consider a particular family $\mathfrak{G}(Z)$ where Z is any set of $\mathfrak{H}^*$. If Y is a set of $\mathfrak{G}(Z)$, it has Z as a factor. Denote by Y^* the intersection of the remaining factors of Y, and by $\mathfrak{G}^*(Z)$ the family of all sets Z^* thus obtained. $\mathfrak{G}(Z)$ being a family of mutually exclusive sets, we easily infer from (13) and (15) that the family $\mathfrak{G}^*(Z)$ is also a family of mutually exclusive sets; furthermore that the correspondence $Y \dashrightarrow Y^*$ between the sets of $\mathfrak{G}(Z)$ and $\mathfrak{G}^*(Z)$ is one-to-one, and that therefore these two families have the same power. On the

other hand, $\mathfrak{G}^*(Z)$ consists of sets of the l^{th} order; thus, by applying to $\mathfrak{G}^*(Z)$ our inductive premise, we obtain:

$$\mathfrak{c}[\mathfrak{G}(Z)] = \mathfrak{c}[\mathfrak{G}^*(Z)] < n \quad \text{for every} \quad Z \in \mathfrak{H}^*. \tag{19}$$

$\mathfrak{n}$ being a regular, and thus an infinite, cardinal number, we also have, by (6) and (17).

$$\mathfrak{c}(\mathfrak{H}^*) \leqq \mathfrak{c}(\mathfrak{H}_{\xi_1}) + \cdots + \mathfrak{c}(\mathfrak{H}_{\xi_{l+1}}) < \mathfrak{n};$$

and hence, in view of (19),

$$\sum_{Z \in \mathfrak{H}^*} \mathfrak{c}[\mathfrak{G}(Z)] < \mathfrak{n}. \tag{20}$$

From (18) and (20) it follows at once that $\mathfrak{G}$ has power $< \mathfrak{n}$. Thus (16) holds for every positive integer l.

We can now extend (16) in the following way:

(21) Every family $\mathfrak{G} \subseteq \mathfrak{F}$ of mutually exclusive sets of any finite orders has power $< \mathfrak{n}$.

In fact, denote by $\mathfrak{G}_l$ the family of those sets $X \in \mathfrak{G}$ which are of the l^{th} order. We obtain the decomposition

$$\mathfrak{G} = \mathfrak{G}_1 \cup \cdots \cup \mathfrak{G}_l \cup \cdots,$$

whence

$$\mathfrak{c}(\mathfrak{G}) \leqq \mathfrak{c}(\mathfrak{G}_1) + \cdots + \mathfrak{c}(\mathfrak{G}_l) + \cdots.$$

(The families $\mathfrak{G}_1, \cdots, \mathfrak{G}_l, \cdots$ are not necessarily mutually exclusive.) On the other hand, we have by (16):

$$C(\mathfrak{G}_l) < \mathfrak{n} \text{ for every positive integer } l.$$

Hence, $\mathfrak{n}$ being by hypothesis a regular number $> \aleph_0$, we easily obtain:

$$\mathfrak{c}(\mathfrak{G}) \leqq \mathfrak{c}(\mathfrak{G}_1) + \cdots + \mathfrak{c}(\mathfrak{G}_l) + \cdots < \mathfrak{n}.$$

Finally we can show that

(22) Every family $\mathfrak{G} \subseteq \mathfrak{F}$ of mutually exclusive sets has 0 power $< \mathfrak{n}$.

For, by (11), every set $Y \in \mathfrak{G}$ is a union of finite intersections of sets $X \in \mathfrak{H}$ and their complements $N - X$. Furthermore, from (5) and (10) it follows that the complement $N - X$ of a set $X \in \mathfrak{H}$ is a union (not necessarily finite) of sets of $\mathfrak{H}$. Therefore the set Y can be represented as a union of finite intersections of sets $X \in \mathfrak{H}$. Hence we can correlate with every set $Y \in \mathfrak{G}$ (which is not empty) a non-empty subset $Y^* \subseteq Y$ of finite order. The family $\mathfrak{G}^*$ of all the sets Y^* thus obtained has clearly the same power as $\mathfrak{G}$, and by (21) this power is $< \mathfrak{n}$.

From the definition of $\mathfrak{d}$, (22) implies:

$$\mathfrak{d}(\mathfrak{G}) \leqq \mathfrak{n};$$

and this formula together with (12) gives:

$$\mathfrak{c}(\mathfrak{F}) = \mathfrak{d}(\mathfrak{F}) = n. \tag{23}$$

Thus our theorem has been proved for a particular set N of power $\mathfrak{n}$. Now consider an arbitrary set S of power $\geqq \mathfrak{n}$. The set S contains a subset N_1 of power $\mathfrak{n}$. We can establish a one-to-one correspondence between the subsets of N and those of N_1, and by means of this correspondence we can construct a field $\mathfrak{F}_1$ of sets $X \subseteq N_1 \subseteq S$ which satisfies (23). This brings the proof to an end.

COROLLARY 1. *If $\mathfrak{n}$ is a regular number $> \aleph_0$, then for every number $\aleph \geqq \mathfrak{n}$ there exists a topological space S of power $\aleph$ such that the family $\mathfrak{G}$ of all open sets (or of all sets which are both open and closed, or of all regular open sets) satisfies the condition: $\mathfrak{d}(G) = \mathfrak{n}$.*

PROOF. Consider any set S of power $\aleph$. By Theorem 2 there is a field $\mathfrak{F}$ of subsets of S such that $\mathfrak{d}(\mathfrak{F}) = \mathfrak{n}$ and that every family $\mathfrak{H} \subseteq \mathfrak{F}$ of mutually exclusive sets has power $< \mathfrak{n}$. We can assume that $S \in \mathfrak{F}$; for otherwise, we could replace $\mathfrak{F}$ by the field $\mathfrak{F}'$ consisting of all the sets $X \in \mathfrak{F}$ and their complements $S - X$, and we could easily show that $\mathfrak{F}'$ still satisfies the conclusion of Theorem 2. Now we correlate with every subset $X \in S$ a new subset $\bar{X} \in S$, the closure of X, by defining $\bar{X}$ as the intersection of all the sets $Y \in \mathfrak{F}$ which contain $X (Y \supseteq X)$. It is easily seen that with this definition S becomes a topological space.

In this space, F is contained in the family $\mathfrak{G}'$ of all those sets which are both open and closed, and the family G of all open sets is constituted by all the unions of the sets $X \in \mathfrak{F}$. Hence it follows that $\mathfrak{d}(G) = \mathfrak{d}(F) = \mathfrak{n}$, and that $\mathfrak{G}$, like $\mathfrak{F}$, does not contain any family $\mathfrak{H}$ of mutually exclusive sets with power $\mathfrak{c}(\mathfrak{H}) = \mathfrak{n}$. Finally it may be noticed that $\mathfrak{G}'$ and the family $\mathfrak{G}''$ of all regular open sets also have these two properties, for we clearly have

$$F \subset \mathfrak{G}' \subset \mathfrak{G}'' \subset \mathfrak{G}.$$

COROLLARY 2. *If $\mathfrak{n}$ is a regular number $> \aleph_0$, then there exists a complete Boolean algebra B such that $\mathfrak{d}(B) = \mathfrak{n}$.*

PROOF. This corollary follows directly from Corollary 1 since, as is well known, the regular open sets of an arbitrary topological space form a complete Boolean algebra,[3] and any two elements of this algebra are disjoint if, and only if, they are disjoint in the usual set-theoretic sense.

Corollary 2 formulates a condition which is necessary for a cardinal number $\mathfrak{n}$ to be regular and $> \aleph_0$. From Theorem 1 it follows that this condition is at the same time a sufficient one. It is easily seen that in this necessary and sufficient condition the term "Complete Boolean algebra" can be replaced by "partially ordered set", "ring (or field) of sets", "family of all open sets of topological space," and so on. If we restrict ourselves to the case of limit numbers, we obtain a necessary and sufficient condition for a number $\mathfrak{n} > \aleph_0$ to be weekly inaccessible.

[3] This result was first stated in A. Tarski, Les fondements de la géometrie de corps, Commemoration of the first Polish Math. Congress, Kracow 1929, p. 42; see also G. Birkhoff op. cit. p. 102.

Finally we give a result of a more special nature:

COROLLARY 3. *The following two sentences are equivalent*:

(*i*) *In every topological space of power* $\leqq 2^{\aleph_0}$ *there exists a family of mutually exclusive open sets with a maximal power.*

(*ii*) *There is no weakly inaccessible number* $\mathfrak{n}$ *which is* $> \aleph_0$ *and* $\leqq 2^{\aleph_0}$.

PROOF. From Corollary 1 it follows immediately that (i) implies (ii); the implication in the opposite direction can be easily derived from Theorem 1.

GENERAL REMARKS ON INACCESSIBLE NUMBERS

In connection with the last corollary it should be noticed that the problem as to whether there exist weakly inaccessible numbers i.e. regular limit numbers which are $> \mathfrak{n}$ and $< 2^{\mathfrak{n}}$ for an infinite number $\mathfrak{n}$ is so far unsolved and probably can not be solved at all within the present systems of general set theory. By definition the weakly inaccessible numbers can not be obtained from smaller ones by such operations as those of infinite addition or of passage from one number to the next greater number. However it is by no means settled that they can not be obtained from smaller ones by means of the other arithmetical operations, namely multiplication and exponentiation. For this reason we single out among the weakly inaccessible numbers a more special class the so called strongly inaccessible numbers, i.e., the numbers which can not be obtained from smaller ones by an arithmetical operation. While e.g. $2^{\aleph_0}$ is clearly not a strongly inaccessible number, it is not known whether this number is weakly inaccessible.

If we enrich the axiom system of set theory by adding the so-called generalized hypothesis of Cantor (which asserts that there is no cardinal number $> \mathfrak{n}$ and $< 2^{\mathfrak{n}}$ for any infinite number $\mathfrak{n}$), we can easily show that the two kinds of inaccessible numbers coincide. However nothing compels us to regard the generalized hypothesis of Cantor as the only possible basis for set-theoretic investigations, and we can equally well consider the possibility of enriching the axioms of set theory by other axioms which contradict the hypothesis of Cantor. For instance it seems quite plausible that the following hypothesis would constitute a consistent and fertile addition to the set theoretical axioms:

Hypothesis of inaccessible numbers: For every infinite number n, $2^{\mathfrak{n}}$ is the smallest weakly inaccessible number $> \mathfrak{n}$.

Furthermore we should like to point out that many set theoretical problems are known at present whose solution involves the notion of an inaccessible number. The first problems of this kind were formulated more than thirty years ago; their number has however considerably increased in recent years. Like the problem solved in the present paper most of these problems can be presented in the following form: Is it true that a certain cardinal number $\mathfrak{n}$ has property P?

We want to give here a few examples of such problems:

PROBLEM 1. (*The representation problem.*) *Is it true that every* $\mathfrak{n}$*-additive and* $\mathfrak{n}$*-distributative Boolean algebra is isomorphic with an* $\mathfrak{n}$*-additive field of sets?* (A Boolean algebra B is called $\mathfrak{n}$-additive if for every set $X \subseteq B$ with $\mathfrak{c}(X) < \mathfrak{n}$

there is an element $y \in B$ such that $y = \bigcup_{x \in X} x$. An $\mathfrak{n}$-additive Boolean algebra is called $\mathfrak{n}$-distributive if

$$\bigcap_{i \in \mathfrak{H}} \bigcup_{j \in \mathfrak{G}_i} x_{i,j} = \bigcup_{f} \bigcap_{i \in \mathfrak{H}} x_{i,f(i)},$$

where $\mathfrak{H}$ is any non-empty set with $\mathfrak{c}(\mathfrak{H}) < \mathfrak{n}$; $\mathfrak{G}_i$ (for $i \in \mathfrak{H}$) are any non-empty sets with $\mathfrak{c}(\mathfrak{G}_i) < \mathfrak{n}$; $x_{i,j}$ is always an element of B, and f runs through all functions which correlate with every element $i \in \mathfrak{H}$ an element $j_i \in \mathfrak{G}_i$. The number of functions f is in general $\geqq \mathfrak{n}$, but it is assumed that the existence of $\bigcap_{i \in H} \bigcup_{j \in j_i} X_{i,j}$ implies that of $\bigcup_f \bigcap_{i \in H} x_{i,f(i)}$.)

PROBLEM 2. (*The prime ideal problem.*) *Is it true that the field of all subsets of a set N with power $\mathfrak{c}(N) = \mathfrak{n}$ contains an $\mathfrak{n}$-additive ideal which is not a principal ideal?* (A family $\mathfrak{F}$ is called n-additive if for every family $\mathfrak{G} \subseteq \mathfrak{F}$ with $\mathfrak{c}(\mathfrak{G}) < \mathfrak{n}$ the union $\bigcup_{X \in \mathfrak{G}} X$ also belongs to $\mathfrak{F}$.)

This problem can also be formulated as that of the existence of an $\mathfrak{n}$-additive non trivial two-valued measure defined over all the subsets of a set N with $\mathfrak{c}(N) = \mathfrak{n}$.

PROBLEM 3. (*The set-function problem.*) *Is it true that there exists an $\mathfrak{n}$-additive and $\mathfrak{n}$-multiplicative set-function defined over all sybsets of a set N of power $\mathfrak{n}$, which is not absolutely additive and absolutely multiplicative?* (By a set function we mean here a function G which correlates with every set X of a certain family $\mathfrak{F}$ another set $G(X)$ which need not belong to the same family. A set function G is called $\mathfrak{n}$-additive or $\mathfrak{n}$-multiplicative, if for every family $\mathfrak{H} \subseteq \mathfrak{F}$ with $\mathfrak{c}(\mathfrak{H}) < \mathfrak{n}$ we have:

$$G\Big(\bigcup_{X \in \mathfrak{H}} X\Big) = \bigcup_{X \in \mathfrak{H}} (G(X)) \quad \text{or} \quad G\Big(\bigcap_{X \in \mathfrak{H}} X\Big) = \bigcap_{X \in \mathfrak{H}} (G(X)),$$

respectively. If these formulas hold for every family $\mathfrak{H} \leqq \mathfrak{F}$, the set function is called absolutely additive or absolutely multiplicative.)

PROBLEM 4. (*The graph problem.*) *Is it true that if a complete graph G of power $\mathfrak{n}$ is split into two graphs G_1 and G_2, at least one of them contains a subgraph of power $\mathfrak{n}$?* (A graph is to be defined as an arbitrary set of non-ordered couples (x, y) with $x \neq y$. By a complete graph of power $\mathfrak{n}$ we mean the set of all such couples formed from the elements of a set N of power $\mathfrak{n}$.)

PROBLEM 5. (*The ordering problem.*) *Is it true that every ordered set N of power $\mathfrak{n}$ contains a subset X of power $\mathfrak{n}$, which is either well ordered, or becomes well ordered if we invert the ordering relation.*

PROBLEM 6. (*Ramification problem.*) *Let ν be the smallest ordinal number such that the power of all ordinals $\xi < \nu$ is $\mathfrak{n}$. Is it true that every ramification system of the ν^{th} order, in which the set of all elements of the ξ^{th} order has power $< \mathfrak{n}$ for every $\xi < \nu$, contains a well-ordered subset of the type ν.* (By a ramification system S we understand a partially ordered set which has the property that, for every $x \in S$, the set $S(x)$ of all elements $y \leqq x$ is well ordered; If the set $S(x)$ is of the type ξ the element x is said to be of the ξ^{th} order. The order of the

whole set S is the smallest ordinal number greater than the order of all elements of S.)

None of these six problems has yet been entirely solved. It can be shown that the solution of these problems is positive for $\mathfrak{n} = \aleph_0$ and is negative for every infinite number $\mathfrak{n} > \aleph_0$ which is not strongly inaccessible, in the case of problems 1–5. (In the case of problem 6 it has been only shown that the solution is negative if $\mathfrak{n}$ is not inaccessible and the generalized Cantor hypothesis holds.) All the problems remain open in case of strongly inaccessible numbers $> \aleph_0$.[4]

This situation is rather typical of the problems involving the notion of an inaccesssible number, which we have here in mind. Most of them so far have resisted all attempts at solution in the case in which $\mathfrak{n}$ is an inaccessible number $> \aleph_0$; it depends, however, on the nature of the problem whether strongly or weakly inaccessible numbers are involved. Ths situation differs slightly in connection with certain problems from the theory of ordered sets. Here the solution is positive for $\aleph_0$, and for all regular numbers which are not weakly inaccessible, and is negative for infinite singular numbers; but the problem again remains open in the case of weakly inaccessible numbers $> \aleph_0$.

The difficulties which we meet in attempting to solve the problems under consideration do not seem to depend essentially on the nature of inaccessible

---

[4] The solution of Problem 1 was given for $\mathfrak{n} = \aleph_0$ by M. H. Stone (see G. Birkhoff op. cit. p. 89.) The solution for numbers which are not inaccessible and $> \aleph_0$ was recently found by A. Tarski and has not yet been published.

For the solution of Problem 2. see A. Tarski, Fund. Math. Vol. 15, p. 42–50. (Une contribution à la théorie de la mesure) (the case $\mathfrak{n} = \aleph_0$). For the case when $\mathfrak{n}$ is not inaccessible, see A. Tarski, Fund. Math. Vol. 30 (1938) p. 150 (Dritter überderkungnatz.)

The solution of Problem 3 for $\mathfrak{n} = \aleph_0$ was given by S. Ulam, Fund. Math. Vol. 16 p. 140–150. (Zur Masstheorie in der allgemeinen Mengenlehre.) The solution for numbers which are not inaccessible follows from a general theorem of A. Tarski; C. R. Soc. Varsovie, Vol. 30 p. 158 (Theorem 2.18).

The solution of Problem 4 was given for $\mathfrak{n} = \aleph_0$ by Ramsay on a problem of formal logic, Proc. London Math. Soc. (2), 30; and for the numbers $\mathfrak{n} > \aleph_0$ which are not inaccessible by P. Erdös, appear in Revista de Tucuman).

The solution of Problem 5 is obvious for $\mathfrak{n} = \aleph_0$; for the numbers $\mathfrak{n} > \aleph_0$ which are not inaccessible the solution was given by Hausdorff, Mengenlehre (1914) p. 145–146. He does not state the solution explicitly, but it can be deduced easily from his results.

The solution of Problem 6 was given for $\mathfrak{n} = \aleph_0$ by D. König, Über eine Schlussweise aus dem endlichen ins unendliche, Acta Szeged, 3, p. 121–130. For the numbers $\mathfrak{n} > \aleph_0$ which are not inaccessible it was given by Aronsajn. It can be shown that the positive solution of Problem 1 for inaccessible numbers $> \aleph_0$ would imply the positive solution of Problem 2; the positive solution of Problem 2 implies that of Problems 3, 4, and 5; also the positive solution of 3 implies that of 2, so that 2 and 3 are equivalent. Further, the positive solution of 4 implies that of 5, and the positive solution of Problem 6 for strongly inaccessible numbers can be deduced from that of Problem 2 (however this solution can also be obtained from weaker hypotheses and can be extended to all inaccessible numbers). Also the positive solution of Problem 6 implies that of Problems 4 and 5. Finally the positive solution of Problem 6 implies the positive solution of Problem 1 in the special case when the Boolean algebra contains only $\mathfrak{n}$ elements. The proof of these equivalences is as yet unpublished.

numbers. In most cases the difficulties seem to arise from lack of devices which enable us to construct maximal sets which are closed under certain infinite operations. It is quite possible that a complete solution of these problems would require new axioms which would differ considerably in their character not only from the usual axioms of set theory, but also from those hypotheses whose inclusion among the axioms has previously been discussed in the literature and mentioned previously in this paper (e.g., the existential axioms which secure the existence of inaccessible numbers, or from hypotheses like that of Cantor which establish arithmetical relations between the cardinal numbers.)

If we now compare the problem which has been actually solved in this paper with those which we have recently discussed, we see that the peculiarity of our problem consists in two facts. First, our problem has been solved for all cardinal numbers, although the inaccessible numbers are essentially involved in the solution. And secondly the number $\aleph_0$ behaves in the discussion of the problem like a singular limit number, and in a directly opposite way to the other regular or inaccessible numbers.

UNIVERSITY OF PENNSYLVANIA AND INSTITUTE FOR ADVANCED STUDY

ANNALS OF MATHEMATICS
Vol. 44, No. 2, April, 1943

ON SOME CONVERGENCE PROPERTIES OF THE INTERPOLATION POLYNOMIALS

P. ERDÖS

(Received June 29, 1942)

It is well known that there exist continuous functions whose Lagrange interpolation polynomials taken at the roots of the Tchebycheff polynomials $T_n(x)$ diverge everywhere in $(-1, +1)$.[1] On the other hand a few years ago S. Bernstein proved the following result[2]: Let $f(x)$ be any continuous function; then to every $c > 0$ there exists a sequence of polynomials $\varphi_n(x)$ where $\varphi_n(x)$ is of degree $n - 1$ and it coincides with $f(x)$ at, at least $n - cn$ roots of $T_n(x)$ and $\varphi_n(x) \to f(x)$ uniformly in $(-1, +1)$.

Fejér proved the following theorem[3]: Let the fundamental points of the interpolation be a normal[4] point group

$$\begin{matrix} & x_1^{(1)} & \\ x_1^{(2)}, & & x_2^{(2)} \\ \cdots & \cdots & \cdots \end{matrix};$$

then for every continuous $f(x)$ there exists a sequence of polynomials $\varphi_n(x)$ of degree $\leqq 2n - 1$ such that $\varphi_n(x_i^{(n)}) = f(x_i^{(n)})$, $i = 1, 2, \cdots n$ and $\varphi_n(x) \to f(x)$ uniformly in $(-1, +1)$. In the present paper we are going to prove the following more general

THEOREM 1. *Let the point group be such that the fundamental functions $l_k^{(n)}(x)$ are uniformly bounded in $(-1, +1)$. Then to every continuous function $f(x)$ and $c > 0$ there exists a sequence of polynomials $\varphi_n(x)$, such that, 1) the degree of $\varphi_n(x)$ is $\leqq n(1 + c)$, 2) $\varphi_n(x_i^{(n)}) = f(x_i^{(n)})$, $i = 1, 2 \cdots n$, 3) $\varphi_n(x) \to f(x)$ uniformly in $(-1, +1)$.*

Theorem 1 generalizes the result of Fejér in two directions; first the point group is more general since it can be shown[5] that the fundamental functions are uniformly bounded for normal point groups, and secondly the degree of $\varphi_n(x)$ is lowered from $2n - 1$ to $n(1 + c)$.

Theorem 1 does not directly generalize the result of S. Bernstein, but we can prove the following

THEOREM 2. *Let the $x_i^{(n)}$ be such that the fundamental functions are uniformly bounded in $(-1, +1)$; then to every continuous function $f(x)$ there exists a sequence of polynomials $\varphi_n(x)$ of degree $\leqq n - 1$ which coincides with $f(x)$ at, at least n-cn points $x_i^{(n)}$ and $\varphi_n(x) \to f(x)$.*

---

[1] G. Grünwald, Annals of Math. Vol. 37, (1936), p. 908–918.

[2] S. Bernstein, Comptes Rendus de l'Acad. des Sciences Vol.

[3] L. Fejér, Amer. Math. Monthly Vol. 41 p. 12.

[4] Ibid.

[5] Fejér proves this only for the so called strongly normal point groups (ibid). The proof for normal point groups is much more complicated and we do not give it here.

We are not going to give a proof of Theorem 2.

The following problem is due to Fejér: Let the $x_i^{(n)}$ be the equidistant abscissae that is $x_i^{(n)} = -1 + \frac{2i-1}{n}$, $i = 1, 2, \cdots n$. The question is, does there exist to every continuous $f(x)$ a sequence of polynomials $\varphi_n(x)$ of degree $< 2n$ such that $\varphi_n(x_i^{(n)}) = f(x_i^{(n)})$ and $\varphi_n(x) \to f(x)$. In other words, does his result proved for the normal point groups also hold for the equidistant point group. We prove the following

THEOREM 3. *To every continuous function $f(x)$ and to every c there exists a sequence of polynomials $\varphi_n(x)$ of degree $\leqq \frac{\pi}{2} n(1+c)$ such that $\varphi_n(x_i^{(n)}) = f(x_i^{(n)})$ and $\varphi_n(x) \to f(x)$ uniformly in $(-1, 1)$, and it can be shown that the constant $\frac{\pi}{2}$ is the best possible.*

Throughout this paper the c's denote absolute constants not necessarily the same. If there is no danger of confusion we will omit the upper index n in $x_i^{(n)}$, $l_k^{(n)}(x)$ etc.

To prove Theorem 1 we need two lemmas.

LEMMA 1. *Let the point group be such that the fundamental functions are uniformly bounded in $(-1, 1)$ and put $\cos \vartheta_i = x_i$, $x_1 < x_2 < \cdots < x_n$, $\vartheta_1 > \vartheta_2 > \cdots > \vartheta_n$; then*

$$\vartheta_i - \vartheta_{i+1} > \frac{c}{n}.$$ [6]

PROOF. Let $|l_i(x)| < D$, $i = 1, 2, \cdots n$. By a well known theorem of S. Bernstein[7] $|d/d\vartheta\, l_i(\cos\vartheta)| \leqq nD$ and since $l_i(x_i) = 1$, $l_i(x_{i\pm1}) = 0$, we have finally

$$\vartheta_i - \vartheta_{i+1} \geqq \frac{1}{Dn}.$$

LEMMA 2. *Let $-1 \leqq y \leqq 1$, $\cos\theta = y$. Then there exists a polynomial $h_y^{(m)}(x)$ of degree $\leqq 2m$ such that $h_y^{(m)}(y) = 1$, $|h_y(x)| \leqq c$, $-1 \leqq x \leqq 1$ and for $\theta - \theta_0 > \frac{A}{m}$*

$$|h_y^{(m)}(\cos\theta_0)| < c_1 \min\left(1, \frac{1}{m^2(\theta-\theta_0)^2}\right).$$

Denote by $X_i^{(m)}$ and X_{i+1}^m the roots of $T_m(x)$ for which $X_i^{(m)} \leqq y \leqq X_{i+1}^{(m)}$. It is easy to see that

---

[6] If the fundamental functions are uniformly bounded we have

$$\frac{c_1}{n} < \vartheta_i - \vartheta_{i+1} < \frac{c_2}{n}.$$

But the upper estimate is not needed here. (Erdös-Turán, Annals of Math. Vol. 39 (1940) p. 706–707.)

[7] S. Bernstein, Belg. Mém. 1912 p. 19.

$$L_i^{(m)}(y) + L_{i+1}^{(m)}(y) \geqq 1,$$ [8]

where $L_i^{(m)}(y)$ denotes the fundamental polynomials belonging to the roots of $T_m(x)$. Without loss of generality we may assume $L_i(y) \geqq \frac{1}{2}$. It is well known that $|L_i^{(m)}(x)| \leqq \sqrt{2}$, $-1 \leqq x \leqq 1$.[9] Thus since $\theta_i - \theta_{i+1} = \frac{\pi}{m}$ ($\cos\theta_i = X_i$) our lemma will be proved if we can show that for $|\theta_i - \theta_0| > \frac{A}{m}$

$$|h_y^{(m)}(\cos\theta_0)| = |L_i^{(m)}(x_0)/L_i^{(m)}(y)|^2 < \frac{c}{A^2}.$$

But

$$|h_y^{(m)}(\cos\theta_0)| \leqq \frac{4}{m^2(\theta_i - \theta_0)^2} < \frac{c}{A^2}.$$

PROOF of Theorem 1. Let $\psi_{n-1}(x)$ be a polynomial of degree n-1 such that

$$|f(x) - \psi_{n-1}(x)| < \epsilon, \quad -1 \leqq x \leqq 1.$$

Put $f(x_i) - \psi_{n-1}(x_i) = \epsilon_i$. Consider the polynomial of degree $\leqq n(1+c)$ such that

$$\varphi_{n-1}(x) = \psi_{n-1}(x) + \sum_{i=1}^{n} \epsilon_i l_i(x) h_{x_i}^{(m)}(x), \qquad m = \left[\frac{cn}{2}\right].$$

Clearly $\varphi_{n-1}(x_i) = f(x_i)$, $i = 1, 2 \cdots n$. We shall prove that $\varphi_{n-1}(x) \to f(x)$ uniformly in $(-1, 1)$. It suffices to show that

$$|g(x)| = \left|\sum_{i=1}^{n} \epsilon_i l_i(x) h_{x_i}^{(m)}(x)\right| < c\epsilon, \qquad -1 \leqq x \leqq 1.$$

Now

$$|g(x)| < c\epsilon \sum_{i=1}^{n} |h_{x_i}^{(m)}(x)| = c\epsilon \sum_{x_i \geqq x} |h_{x_i}^{(m)}(x)| + c\epsilon \sum_{x_i \leqq x} |h_{x_i}^{(m)}(x)| = c\epsilon\,(\Sigma_1 + \Sigma_2)$$

Thus we only have to show that $\Sigma_1 + \Sigma_2 < c_1$. By Lemma 1,

$$\Sigma_1 < \sum_r |h_{x+k_r}^{m}(x)|,$$

where | are $\cos(x + k_r)$ − are $\cos x$ | > $(rc)/n$. Thus by Lemma 2

$$\Sigma_1 < \sum_r \frac{c_3}{r^2 c^2} < c_4.$$

Similarly we obtain $\Sigma_2 < c_2$, which completes the proof of Theorem 1

[8] Erdös-Turán, Annals of Math. Vol. 41, (1941) p. 529, Lemma IV.

[9] L. Fejér, Mathematische Annalen, Vol. 106, (1932) p. 5.

Theorem 1 does not give a necessary and sufficient condition for the existence of a sequence of polynomials $\varphi_n(x)$ of degree $\leqq n(1+c)$ with $\varphi_n(x_i) = f(x_i)$ and $\varphi_n(x) \to f(x)$ uniformly in $(-1, 1)$. To obtain such a condition let $x_i^{(n)}$ be a point group, put $\cos \vartheta_i^{(n)} = x_i^{(n)}$ and denote by $N_n(a, b)$ the number of the ϑ_i in (a, b), We have the following:

THEOREM 4. *A necessary and sufficient condition that to every continuous function $f(x)$ and to every $c > 0$ there exists a sequence of polynomials $\varphi_n(x)$ of degree $\leqq n(1+c)$ such that $\varphi_n(x_i) = f(x_i)$ and $\varphi_n(x) \to f(x)$ uniformly in $(-1, 1)$ is that if $n(b_n - a_n) \to \infty$, $0 \leqq a_n < b_n \leqq \pi$*

$$(1)\quad \limsup \frac{N_n(a_n, b_n)}{n(b_n - a_n)} \leqq \frac{1}{\pi} \quad \text{and} \quad \liminf (\vartheta_i - \vartheta_{i+1})n > 0, \ (n \to \infty \ i \text{ arbitrary})$$

Condition (1) *states that the number of ϑ_i in (a_n, b_n) can not be much greater than the number of roots of $T_n(x)$ in (a_n, b_n). If the fundamental functions $l_k(x)$ are uniformly bounded* (1) *is satisfied, for then we have*

$$\lim \frac{N_n(a_n, b_n)}{n(b_n - a_n)} = \frac{1}{\pi}, \qquad n(b_n - a_n) \to \infty$$ [10]

We do not give the proof of Theorem 4, but the following proof of Theorem 3 can by a simple modification be applied to it.

PROOF of Theorem 3. Here the fundamental points are

$$x_i^{(n)} = -1 + \frac{2i-1}{n}$$

First we prove the existence for every n and $c > 0$ of $m = \frac{\pi}{2} n(1+c)$ points. $y_i^{(m)}$, $i = 1, 2, \cdots n$ such that (I) the $x_i^{(n)}$ occur among the $y_i^{(m)}$ (II) the fundamental functions $L_k(x)$, $k = 1, 2, \cdots m$ are uniformly bounded in $(-1, 1)$ (The $L_k(x)$ are the fundamental functions belonging to the $y_i^{(m)}$). Having constructed the $y_i^{(m)}$ satisfying (I) and (II) we immediately obtain Theorem 3 by applying Theorem 1.

To construct the $y_i^{(m)}$ we first remark that by putting

$$\cos \vartheta_i = -1 + \frac{2i-1}{n}, \qquad i = 1, 2, \cdots n$$

we obtain by a simple calculation

$$\vartheta_i - \vartheta_{i+1} > \frac{\pi}{m}$$

Now we construct a sequence $y_i^{(m)}$, $i = 1, 2, \cdots m$ such that (1) the $x_i^{(n)}$ occur among the $y_i^{(m)}$ (2) put $\cos \theta_i = y_i$ then $\theta_i = \frac{2i-1}{m} \frac{\pi}{2} + \frac{d_i}{m}$ where $\sum_{i=1}^{k} d_i$ is uniformly bounded (3) $\theta_i - \theta_{i+1} \geqq \frac{\pi}{4m}$. (2) and (3) insure that the y_i are "very nearly" the roots of $T_m(x)$.

[10] Erdös-Turán, ibid., p. 519.

We construct the y_i as follows: Suppose $y_i < y_2 < \cdots < y_{i-1}$ are already constructed. We further make the hypothesis that if ϑ_r ($\cos \vartheta_r = x_r$) is the greatest $\vartheta < \theta_{i-1}$ then $\theta_{i-1} - \vartheta_r > \frac{\pi}{4m}$. If $\sum_{j=1}^{i-1} d_j < 0$ we choose for y_i either the least $x_r > y_{i-1}$, or if $\vartheta_r < \theta_{i-1} - \frac{4\pi}{m}$ we put $\theta_i = \theta_{i-1} - \frac{2\pi}{m}$. Thus θ_i does not come nearer than $\frac{\pi}{4m}$ to the greatest $\vartheta < \theta_i$. If $\sum_{j=1}^{i-1} d_j > 0$, $y_i = x_r$ if $\vartheta_r > \theta_{i-1} - \frac{\pi}{2m}$ and $\theta_{i-1} - \frac{\pi}{4m}$ otherwise. Thus in any case if ϑ_j is the greatest $\vartheta < \theta_i$ then $\theta_i - \vartheta_j > \frac{\pi}{4m}$. In this we can construct $y_1, y_2, \cdots y_m$. (1) and (3) are clearly satisfied and it is quite immediate that (2) is also satisfied. Now we have to show that the y_i's satisfying (1), (2), and (3) also satisfy (I) and (II). (I) is clearly satisfied, the proof that (II) is satisfied is slightly more difficult. Denote by $z_1, z_2, \cdots z_m$ the roots of $T_m(x)$ and by $L_k'(x)$ the fundamental functions belonging to the z_i. From (2) and (3) it follows by a simple calculation that

$$c_1 \omega'(y_k) < T_m'(z_k) < c_2 \omega'(y_k), \qquad \omega(x) = \prod_{i=1}^{m} (x - y_i) \tag{2}$$

where c_1 and c_2 are independent of m and k. Denote

$$\max_{-1 \leq x \leq 1} |L_k(x)| = A_k, \qquad \max_{-1 \leq x \leq 1} L_k'(x) = B_k$$

Then again from (2) and (3) by a simple calculation [using (2)]

$$c_3 > \frac{A_k}{B_k} < c_4. \tag{3}$$

We know that[11]

$$B_k < \sqrt{2}. \tag{4}$$

Thus from (3) and (4) we obtain (3), and this completes the proof of Theorem 3.

To obtain the second part of Theorem 3 we first have to prove

LEMMA 3. *Let $m = [(\pi/2)n(1 - \epsilon)]$, $\epsilon > 0$ fixed, independent of m and n, n odd. Let $\varphi_m(x)$ be a polynomial of degree m such that $\varphi_m(0) = 1$ and $\varphi_m(-1 + ((2i - 1)/n)) = 0$, $i = 1, 2, \cdots [(n - 1)/2], [(n + 3)/2] \cdots n$.* Then

$$\max_{-1 \leq x \leq 1} |\varphi_m(x)| > c_1^n, \qquad c_1 = c_1(\epsilon) > 1.$$

PROOF. We use the following lemma due to M. Riesz[12]: Let $\varphi_m(x)$ be a poly-

[11] L. Fejér, see footnote 9.

[12] M. Riesz, Jahresbericht der Deutschen Math. Vereinigung, (1915), p. 354–368.

nomial of degree m, it assumes its absolute maximum in $(-1, 1)$ at the point $x_0 = \cos \vartheta_0$. Let $x_i = \cos \vartheta_i$ be the nearest root of $\varphi_m(x)$ in $(-1, +1)$ then

$$|\vartheta_i - \vartheta_0| \geqq \frac{\pi}{2m}.$$

It immediately follows from this lemma that if x_i and x_{i+1} are the nearest roots including x_0, then we have

$$\vartheta_i - \vartheta_{i+1} \geqq \frac{\pi}{m}.$$

Put now $\cos \vartheta_i = -1 + (2i - 1)/n$. A simple calculation shows that there exists a constant $c_2 = c_2(\epsilon)$ such that if $-c_2 \leqq x_i < x_{i+1} \leqq c_2$ then

$$\vartheta_i - \vartheta_{i+1} < \frac{\left(1 - \frac{\epsilon}{2}\right)\pi}{m}.$$

Hence $\varphi_m(x)$ can not assume its absolute maximum for $-c_2 \leqq x \leqq c_2$ except if

$$\frac{x_{n-1}}{2} < x < \frac{x_{n+3}}{2} \qquad \text{(i. e. in the neighborhood of 0)}$$

Consider now a polynomial $h_m(x)$ with highest coefficient the same as that of $\varphi_m(x)$ whose roots are defined as follows: Let $-c_2 < z_i < c_2$ then $z_i = (1 + \delta)x_i$ where δ is chosen so small that

$$\theta_i - \theta_{i+1} < \frac{1 - \frac{\epsilon}{4}}{m}\pi \qquad (\cos \theta_i = z_i)$$

The other roots of $h_m(x)$ coincide with those of $\varphi_m(x)$. Clearly the degree of $h_m(x)$ is m. Define

$$g(x) = \left(x + \frac{1}{4m}\right)\left(x - \frac{1}{4m}\right)h_m(x).$$

By the lemma of M. Riesz $g(x)$ does not assume its absolute maximum in $(-c_2, c_2)$. It follows from the inequality of the arithmetic and geometric means that

$$|g(x)| < |\varphi_m(x)| \text{ for } c_2(1 + \delta) \leqq |x| \leqq 1 \tag{5}$$

Denote by $A(c_2)$ the number of x_i in $(-c_2, +c_2)$. We evidently have $A(c_2) > c_3 n$. Thus

$$|g(0)| > \varphi_m(0)|(1 + \delta)^{c_3 n}\frac{1}{16m^2} > \varphi_m(0)c_1^n = c_1^n (c_1 > 1) \tag{6}$$

But since $g(x)$ assumes its absolute maximum in $(-1, 1)$ for some $|x_0| > c_2(1 + \delta)$ we have by (5) and (6)

$$|\varphi_m(x_0)| > |g(x_0)| > c_1^n \quad \text{q.e.d.}$$

Let now $n_1, n_2, \cdots$ be an infinite sequence of odd integers, which tend to infinity sufficiently quickly. We define a polynomial $\psi_i(x)$ as follows

$$\psi_i\left(-1+\frac{2j-1}{n_r}\right)=0, \qquad r \leqq i, \qquad j \neq \frac{1+n_r}{2}, \qquad j \leqq n_r,$$

$$\psi_i(0)=1, \qquad |\psi_i(x)| \leqq 2, \qquad -1 \leqq x \leqq 1.$$

From the approximation theorem of Weierstrass it follows that such a $\psi_i(x)$ exists. Consider now the continuous function

$$f(x)=\sum_{k=1}^{\infty}\frac{\psi_k(x)}{2^k}.$$

If the second part of Theorem 3 would not be true, we could find a sequence of polynomials $\varphi_i(x)$ of degree $\leqq n_i(\pi/2)(1-\epsilon)$ such that $\varphi_i(-1+((2j-1)/n_i)) = f[1+((2j-1)/n_i)]$ and $\varphi_i(x) \to f(x)$ uniformly in $(-1, 1)$. For $k > i$

$$\psi_k\left(-1+\frac{2j-1}{n_i}\right)=0, \qquad j \neq \frac{1+n_i}{2}.$$

Thus $\varphi_i(x)$ coincides with

$$\sum_{r=1}^{i-1}\frac{\psi_r(x)}{2^r}=g(x)$$

at the points $-1+((2j-1)/n_i)$, $j \neq ((1+n_i)/2)$.

Let now n_i tend to infinity so quickly that n_i is greater than the degree of $g(x)$. Then $\varphi_i(x)$ can be written as

$$\varphi_i(x)=\varphi_i^{(1)}(x)+\varphi_i^{(2)}(x),$$

where $\varphi_i^{(1)} = g(x)$, and $\varphi_i^{(2)}(x)$ is of degree $\leqq ((\pi/2)-c_1)n_2$ and $\varphi_i''(-1+((2j-1)/n_i))=0, j \neq ((1+n_i)/2), j \leqq n_i$, also $\varphi_i^{(2)}(0)=\sum_{k \geqq i}((\psi_k(0))/2^k) = (1/2^{i-1})$. Thus by lemma 3

$$\max_{-1 \leqq x \leqq 1}|\varphi_i(x)| \geqq \max_{-1 \leqq x \leqq 1}|\varphi_i^{(2)}(x)| - 2 > \frac{c_2^{n_i}}{2^{i-1}} > c_3^{n_i} \ (c_2 \text{ and } c_3 \text{ are} > 1$$

if n_i tends to infinity sufficiently quickly. Hence $\varphi_i(x)$ can not converge uniformly to $f(x)$, and this completes the proof of Theorem 3.

By a more complicated argument we could prove that a point x_0 exists such that $\varphi_n(x_0)$ diverges. We give only the sketch of the proof. Since $\max_{-1 \leqq x \leqq 1}|\varphi_{n_i}(x)| > (1+\delta)^{n_i}$ it follows from a theorem of Remes[13] that there exists in $(-1, 1)$ a set of measure $> c = c(\delta)$ such that on this set $|\varphi_{n_i}(x)| > (1+(\delta/2))^{n_i}$. Then, it follows easily that there exist a point x_0 with $\limsup |\varphi_{n_i}(x_0)| = \infty$.

[13] E. Remes, *Sur une propriété extremale des polynomes de Tchebycheff*. Comm. de l'Institut des Sciences etc. Kharkov, (1936) série 4, XIII fasc. 1, p. 93–95.

By the same method we can prove the following:

THEOREM 5. *Let* $x_1^{(1)}$ / $x_2^{(1)}$, $x_2^{(2)}$ / $\cdots$ *be a point group and put* $\cos(\vartheta_i^{(n)}) = x_i^{(n)}$. *Suppose that*

$$\liminf n(\vartheta_i^{(n)} - \vartheta_{i+1}^{(n)}) = \frac{\pi}{d}, \quad (n \to \infty \text{ , } i \text{ arbitrary})$$

Then to every continuous $f(x)$ *and constant* $c > 0$ *there exists a sequence of polynomials* $\varphi_n(x)$ *of degree* $< d(1 + c)n$ *such that* $\varphi_n(x_i^{(n)}) = f(x_i^{(n)})$ *and* $\varphi_n(x) \to f(x)$ *uniformly in* $(-1, 1)$.

The constant d_1 of Theorem 5 is not best possible. We can obtain the best possible constant d_1 as follows: Let a_n and b_n be two arbitrary sequences of real numbers, such that $0 \leqq a_n < b_n \leqq \pi$, $n(b_n - a_n) \to \infty$. Then if $d < \infty$

$$\limsup \frac{N_n(a_n, b_n)}{n(b_n - a_n)} = \pi d_1 .$$

Lemma 3 would not suffice for the proof of Theorem 5. Here we need

LEMMA 4. *Let* $\varphi_n(x)$ *be a polynomial of degree* n, $\varphi_n(0) = 1$. *Let* $\psi(n)$ *be any function of* n *tending to infinity together with* n *and let* c_1 *be a constant independent of* n. *Then if* $\varphi_n(x)$ *is such that for every* $c_1 < A < \psi(n)$ *the number of roots of* $\varphi_n(\cos\vartheta)$ *in* $((\pi/2) - (A/n), (\pi/2) + (A/n))$ *is greater than* $[((1 + c_2)2)/\pi]$ *we have* $\max_{-1 \leqq x \leqq 1} |\varphi_n(x)| \to \infty$. *Our condition means that the number of roots of* $\varphi_n(x)$ *in the neighborhood of* O *is substantially larger than the number of roots of* $T_n(x)$. *The proof of Lemma 4 is similar, but more complicated than the proof of Lemma 3.*

UNIVERSITY OF PENNSYLVANIA